

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

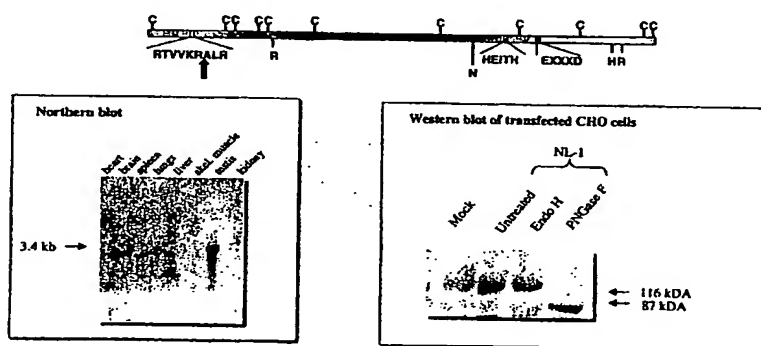


INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁷ : C12N 15/57, 9/64, 15/85, 5/10, C12Q 1/68, C12N 15/62, C07K 16/40, G01N 33/573</p>	<p>A2</p>	<p>(11) International Publication Number: WO 00/47750 (43) International Publication Date: 17 August 2000 (17.08.00)</p>
<p>(21) International Application Number: PCT/CA00/00147 (22) International Filing Date: 11 February 2000 (11.02.00) (30) Priority Data: 2,260,376 11 February 1999 (11.02.99) CA (71) Applicant (for all designated States except US): UNIVERSITE DE MONTREAL [CA/CA]; 2900, Boulevard Edouard-Montpetit, Montreal, Québec H3C 3J7 (CA). (72) Inventors; and (75) Inventors/Applicants (for US only): DESGROSEILLERS, Luc [CA/CA]; 18 des Geais Bleus, St-Basile-Le-Grand, Québec J3N 1L1 (CA). BOILEAU, Guy [CA/CA]; 7645 Malherbe, Brossard, Québec J4Y 1E6 (CA). (74) Agents: DUBUC, Jean, H. et al.; Goudreau Gage Dubuc, The Stock Exchange Tower, Suite 3400, 800 Place Victoria, Montreal, Quebec H4Z 1E9 (CA).</p>		<p>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published Without international search report and to be republished upon receipt of that report.</p>

(54) Title: NEW METALLOPROTEASES OF THE NEPRILYSIN FAMILY

Structure and expression of NL-1



(57) Abstract

In this paper, we describe RT-PCR strategies that allowed us to identify and clone members of the NEP-like family. Degenerate oligonucleotide primers corresponding to consensus sequences located on either side of the HEXXH consensus sequence for zincins were designed and used in RT-PCR with mouse and human testis cDNAs. DNA fragments with lengths expected from the sequence of this class of enzymes were obtained. These DNA fragments were cloned and sequenced. Using this PCR strategy and the PCR fragments as probes to screen cDNA libraries, three zincin-like peptidases were identified in addition of known members of the family. The cDNA sequences allowed to derive specific probes for Northern and *in situ* hybridization, and probe human chromosomes to localize the gene and establish potential links to genetic diseases. Furthermore, these cDNA sequences were used to produce recombinant fusion proteins in *Escherichia coli* in order to raise specific antibodies. Finally, the cDNA sequences were cloned in mammalian expression vectors and transfected in various mammalian cell lines to produce active recombinant enzymes suitable for testing specific inhibitors.

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TITLE OF THE INVENTION

New Metalloproteases of the Neprilysin Family

BACKGROUND OF THE INVENTION

5 Peptides are used by cells from yeast to mammals to elicit physiological responses. The use of peptides as messengers usually involves the following steps: 1) production and release of the peptide by a specific cell, 2) interaction of the peptide with a receptor on the surface of the target cell, and 3) degradation of the peptide to terminate its action. The first and last steps of this scheme require the participation of

10 proteases/peptidases. There is increasing evidence that membrane-associated zinc-metallopeptidases play important roles in both of these steps. Although activation of prohormone precursors into bioactive peptides is generally performed by proteases of the subtilisin family located in the *Trans*-Golgi Network or in secretory granules of the cell (for a review see: (Seidah and Chrétien, 1995)) a few peptides need a final

15 processing step. This step involves the action of membrane-associated zinc-metallopeptidases. Two cases are particularly well documented: angiotensin-converting enzyme (ACE) which cleaves inactive angiotensin I into angiotensin II (Corvol and Williams, 1997) and endothelin-converting enzymes (ECEs) which cleave isoforms of big endothelins into endothelins (Turner, 1997a). In addition to their role

20 in peptide activation, cell surface zinc-metallopeptidases have also been implicated in the termination of the peptidergic signal by breaking down the active peptides into inactive fragments. One of the best known of these peptidases is probably Neutral Endopeptidase-24.11 (Neprilysin, NEP) that has been implicated in the physiological degradation of several bioactive peptides (Kenny, 1993). Interestingly, NEP and the

25 ECEs show significant structural similarities and appear to be members of a family of peptidases that also includes PEX, a newly discovered and not yet characterized peptidase, and the KELL blood group protein (Turner and Tanzawa, 1997b). Because of their important role as regulators of bioactive peptide activity, these enzymes (more specifically NEP and the ECEs) have been identified as putative targets for therapeutic

30 intervention, similar to the way ACE inhibitors are used to control blood pressure. The recent discovery of PEX, another member of the family, which appears to be involved in phosphate homeostasis, raised the possibility that other yet unknown members might exist.

35 Members of the NEP-like family are type II membrane proteins consisting of three distinct domains: a short NH₂-terminal cytosolic sequence, a single transmembrane region, and a large extracellular or ectodomain responsible for the catalytic activity of the enzyme. There are potential N-glycosylation sites and cysteine residues that are involved in disulfide bridges stabilizing the conformation of the active enzyme. These enzymes are metalloenzymes with a Zn atom in their active site. As

such, they belong to the zincin family of peptidases which is characterized by the active site consensus sequence HEXXH (Hooper, 1994), where the two histidine residues are zinc ligands. In members of the NEP-like family of peptidases, the third zinc ligand is a glutamic acid residue located on the carboxy-terminus side of the consensus sequence. This characteristic puts them in the gluzincin sub-family (Hooper, 1994). The model enzyme for gluzincins is thermolysin (TLN) a bacterial protease whose 3D structure has been determined by X-ray crystallography (Holmes and Matthews, 1982). The active site of NEP has been extensively studied by site-directed mutagenesis and several residues involved in zinc binding (Devault et al., 1988b; Le Moual et al., 1991; Le Moual et al., 1994), catalysis (Devault et al., 1988a; Dion et al., 1993), or substrate binding (Vijayaraghavan et al., 1990; Beaumont et al., 1991; Dion et al., 1995; Marie-Claire et al., 1997) have been identified (for a recent review see Crine et al., 1997).

15 SUMMARY OF THE INVENTION

Here, we developed an RT-PCR strategy to look for other members of this important family of peptidases. This strategy allowed the molecular cloning and characterization of three additional NEP-like (NL) metallopeptidases (called NL-1, NL-2 and NL-3). Knowledge obtained through these studies allows the generation of reagents (nucleic acid probes and primers, antibodies and active recombinant enzymes) for further biochemical characterization of these enzymes and their pattern of expression and will greatly help the rational design of specific inhibitors that could be used as therapeutic agents.

Accordingly, the present invention relates to the following products:

- 25 A. Degenerate primers for screening new NEP-related enzymes;
- B. NL-1, NL-2 and NL-3 proteins as NEP-related enzymes;
- C. Nucleic acids encoding these enzymes;
- D. Antibodies directed against the enzymes;
- E. Recombinant vectors comprising the nucleic acids encoding the enzymes and hosts transformed therewith;
- 30 F. Fragments of the nucleic acids useful as probes or primers to hybridize and detect the presence of an NL-1, NL-2 and NL-3 genes, or to hybridize and amplify and produce gene fragments;
- G. Soluble forms of NL-1, NL-2 and NL-3; and
- 35 H. Nucleic acids comprising the N-terminal part of NL-1 or NL-2 which terminates with a sequence encoding a furin recognition site, such nucleic acids being useful for making a fusion protein with the ectodomain of any protein of interest, and for releasing a soluble form of that protein of interest (containing the ectodomain) in the medium.

Also the present invention relates to the following methods:

- A. A method for screening NEP-related enzymes that make use of degenerate primers or probes selected from a region of NEP family members in a highly conserved region, namely around the zinc-binding sites; and
- 5 B. A method for producing NL-1, NL-2 or NL-3 that includes the steps of culturing the above recombinant host and recovering NL-1, NL-2 and NL-3 gene products therefrom.

The present invention will be described hereinbelow by referring to specific
10 embodiments and appended figures, which purpose is to illustrate the invention rather than to limit its scope.

In the first section, general procedures leading to the identification and localization of NL-1, NL-2 and NL-3 are given. In the second section, slightly different procedures are given for completing or reiterating the work performed on NL-1.

15

BRIEF DESCRIPTION OF THE FIGURES

Figure 1: Amino acid sequence comparison of human NEP, PEX, KELL and ECE1 peptidases. Amino acid sequences in boxes are those used to design the oligonucleotide primers. Numbers and arrows under the sequences identify the primer
20 and its orientation.

Figure 2: Sequences of the oligonucleotide primers used in the PCR reactions.

Figure 3: Nucleotide and amino acid sequence of the mouse NL-1 cDNA. The sequence of the DNA fragment obtained by PCR is in brackets.

Figure 4: Partial nucleotide and amino acid sequence of the human NL-2
25 cDNA. The sequence of the DNA fragment obtained by PCR is in brackets.

Figure 5: Partial nucleotide and amino acid sequence of the human NL-3 cDNA.

Figure 6: Amino acid sequence comparison of NEP, NL-1, NL-2 and NL-3 peptidases.

Figure 7: *In situ* hybridization of mouse testis sections using NL-1 as a probe.

Figure 8: *In situ* hybridization of mouse sections using mouse NL-3 as a probe.

Figure 9: *In situ* hybridization of mouse spinal chord sections

Figure 10: Expression of NL-1 in mammalian cells.

Figure 11: Activity of recombinant soluble NL-1.

Figure 12: Expression of a soluble form of NL-3 using NL-1 amino-terminal
35 domain.

DESCRIPTION OF THE PREFERRED EMBODIMENTS OF THE INVENTION

SECTION 1)

5 MATERIALS AND METHODS

DNA and RNA manipulations

All DNA manipulations and Northern blot analysis were performed according to standard protocols (Ausubel et al., 1988; Sambrook et al., 1989).

mRNA purification and cDNA synthesis

10 mRNAs were prepared from mouse testis using Quick Prep Micro mRNA purification kit (Pharmacia Biotech). Purified mRNAs were kept at -70° until ready used. First strand cDNA was synthesized from 1 µg of mRNA using the First-Strand cDNA synthesis kit (Pharmacia Biotech). The human testis cDNA library was obtained from Clontech.

15 Polymerase chain reaction protocol

PCR was performed in a DNA thermal cycler with 5 µl of cDNA template and 1 µl of Taq DNA polymerase in a final volume of 100 µl, containing 1 mM MgCl₂, 2 µM of each primer oligonucleotide, 20 µM of each dNTP and 5% DMSO. Cycling profiles included an initial denaturation step of 5 min at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 40°C and 1.5 min at 72°C. A final extension step was performed at 20 72°C for 10 min. The amplified DNA was loaded on a 2% agarose gel and visualized by staining with ethidium bromide. Fragments ranging in size between 500-700 bp were cut and eluted from the gel. If needed, a second round of PCR was done with nested oligonucleotide primers, using 10 µl of the first PCR reaction, or of the eluted band cut from the agarose gel. Resulting fragments were ligated in pCR2.1 vector (Invitrogen) according to the distributor's recommendations. DH5α *Escherichia coli* cells were transformed with the ligation mixture and grown on 2YT plates in the presence of kanamycin. Plasmids were prepared from resistant cells and sequenced.

25 In situ hybridization on mouse tissues and chromosomal localization of human genes

30 *In situ* hybridization on whole mouse slices or isolated tissues was performed as described previously (Ruchon et al., 1998).

To determine the chromosomal localization of human NL-2 and NL-3 genes, a technique for mapping genes directly to banded human chromosomes was used. Metaphase chromosomes were obtained from lymphocytes cultured from normal human peripheral blood. Cells were synchronized with thymidine and treated with 5-bromodeoxyuridine (BrdU) during the last part of the S phase to produce R-banding. Biotin-labeling of the probe was done by nick-translation (Bionick, BRL) and the probe was visualized by indirect immunofluorescence.

35 Antibody production

To raise antibodies against the new peptidases, the cDNA sequences of each protein was compared to that of other members of the family and the sequence segment showing the less homology was used. These sequences are from amino acid residues 273 to 354 for NL-1, from 75 to 209 for NL-2 and from 143 to 465 for NL-3. These cDNA fragments were cloned in vector pGEX2T (Pharmacia Biotechnology) downstream from and in phase with Gluthatione-S-transferase (GST). Plasmids were transformed in *E. coli* strain AP401 and, induction of synthesis and purification of the fusion proteins were performed as recommended by the supplier. The NL polypeptides were cleaved from the fusion protein with thrombin and purified by SDS-PAGE. NL polypeptides were injected to rabbits or mice according to the following schedules: for rabbits, initial injection of 150 µg of protein with boosts of the same amount 4 weeks and 8 weeks following the initial injection; for mice, initial injection of 100 µg of protein followed by boosts of the same amounts 3 and 6 weeks later. A month after the last injection, sera were collected from the animals and tested by immunoblotting against the initial *E. coli*-produced antigens and the recombinant proteins produced in mammalian cell lines.

Production of monoclonal antibodies

cDNA fragments corresponding to amino acids segments of NLs selected to raise antibodies were used to construct a GST-fusion protein in *E. coli*. This fusion protein was purified from *E. coli* extracts by affinity chromatography on a glutathione-Sepharose column according to the supplier's instructions (Amersham-Pharmacia). After thrombin cleavage, the NL portion of the GST fusion protein was further purified by electroelution from a polyacrylamide gel. This material was used to immunise 4 mice (5 injections of ≈50 µg of NL polypeptide). Blood was collected from each mice after the immunisation schedule and the presence of antibodies in mice serum was assessed by ELISA using microtiter plates coated with NL polypeptide from *E. coli* extracts. Mice sera were also tested for the presence of NL antibodies by Western blotting extracts of mammalian cells transfected with the NL expression vectors. One mouse selected for its high titer of NL specific antibodies (as measured by ELISA) was sacrificed and its spleen cells were collected and immortalised by fusion with myeloma cells (strain: P3-X63Ag.653 from ATCC) as described previously (Crine 1985). Hybridoma cells were selected for their ability to grow in HAT selection medium and cloned by several rounds of limiting dilution. Hybridomas showing proper affinity and specificity to the enzymes NL-1, NL-2 and NL-3 where selected.

Expression of NLs in cultured mammalian cells and enzymatic assays

The cDNAs for NL-1 and NL-3 were cloned in vectors pcDNA3 or pRcCMV (Invitrogen) and introduced by transfection in mammalian cell lines according to procedures already described in our laboratory (Devault *et al.*, 1988a). Procedures to

prepare extracts of cellular proteins or culture media were also described in previous papers (Devault *et al.*, 1988a; Lemay *et al.*, 1989). The presence of NLs in these extracts was monitored by immunoblotting using specific antibodies.

5 Extracts of cellular proteins and culture media were assayed for enzymatic activity. Two tests were performed. The first used [³H]-Tyr-(D)Ala₂-Leu-enkephalin as substrate and was performed according to Lemay *et al.*, (1989). The second used bradykinin as substrate and was performed as described by Raut *et al.* (1999).

RESULTS

10 Cloning of NL-1, a new member of the NEP family

The molecular cloning in the past few years of ECEs, PEX and KELL showed that all these proteins have between 50 and 60% similarity with NEP. This observation led us to believe that these peptidases are part of an extended family and that there could be still additional members to be discovered. To test this hypothesis, we aligned
15 the amino acid sequences of the members of the NEP-like family and designed degenerate oligonucleotide primers to be used in RT-PCR reactions (Figure 1 and 2). These primers were located on either side of the HEXXH consensus sequence for zincins. Because they are highly degenerate, primers 1 and 2 were each subdivided into two pools, 1A-1B, and 2A-2B, respectively (Figure 2). Any PCR amplified DNA
20 fragment that corresponds to a peptidase of the family should normally contain the consensus sequence and be easily recognized by sequencing of the cloned fragments. Using this strategy, we first performed PCR reactions with primer pairs 1A-3 and 1B-3. The amplified DNA migrates mostly as a smear starting at around 700 bp and going down to 100 bp. As the expected fragments should be around 550 bp, we isolated from
25 the gel the section corresponding to DNA fragments longer than 500 bp. A second round of PCR reactions was performed with both crude PCR products of the first reaction and isolated DNA bands, using primers 2A-3 and 2B-3. The expected 296 bp fragment was seen on the gel (not shown).

Cloning of these DNA fragments generated approximately 350 clones, of which
30 44 were sequenced. Nine of these had no inserts or corresponded to sequences not related to the NEP family, 24 corresponded to NEP, 3 to PEX, and 8 corresponded to one putative new member of the family, since they all contained the HEXXH consensus sequence for zincins and showed 65% homology with mouse NEP (in boxes Figure 3). This fragment was then used to screen a mouse testis cDNA library,
35 and allowed us to isolate a complete cDNA of 2592 nucleotides (Figure 3). The identity of this sequence with other members of the family is presented in Table I. This new member was called NL-1, for NEP-like peptidase 1.

Cloning of NL-2 and NL-3.

A strategy similar to that described for amplification of enzymes of the NEP family from mouse testis cDNAs was used with a human testis cDNA library using two different oligonucleotide primers. This time, DNA fragments of approximately 900 bp
5 were obtained and cloned. Ten clones were sequenced, revealing the presence of NEP and two new peptidases of the family that we have called NL-2 and NL-3.

The NL-2 PCR fragment was 879 nucleotides in length and encoded a 293 amino acid residue segment probably located in the carboxy-terminal domain of this putative peptidase (in brackets Figure 4). This PCR fragment was then used to screen
10 a lambda gt10 human brain cDNA library. It allowed the isolation of other cDNA fragments which overlap partially with the NL-2 PCR fragment. Fusion of these lambda clones and the PCR fragment resulted in an open reading frame of 770 amino acid residues. The use of 5' RACE protocols with human testis cDNA libraries allowed completion of the sequence of NL-2 ORF (Figure 4). This ORF codes for a putative
15 protein that is about 80% identical to the mouse NL-1 protein (Figure 6). Across species, members of the NEP, PEX, ECEs sub-families have highly conserved sequences (more than 94% identity). Although a sequence identity of about 80% only exists between the novel human protein and mouse NL-1, these proteins share unique characteristics that make possible the fact that NL-2 protein may be the human
20 homologue of NL-1. The identity of NL-2 with other members of the family is presented in Table I.

The 879 bp PCR fragment encoding NL-3 showed an open reading frame of 293 amino acid residues (Figure 5, in brackets). Sequence analysis of NL-3 showed that it was 94.2 % identical to an EST sequence from mouse embryonic tissue present
25 in publicly accessible DNA data banks. This mouse EST sequence, commercially available from American Tissue and Cell Culture (ATCC), had been obtained previously by our laboratories.

Since Northern blot analysis of human tissues with the NL-3 PCR fragment showed the expression of this protein in spinal chord (see below), the same PCR DNA
30 fragment was used to screen by hybridization a human spinal chord cDNA library constructed in phage λ vectors. One clone contained a full-length ORF of 752 amino acid residues that encompassed the 293 amino acid residue ORF of the PCR fragment. Further probing, cloning and sequencing lead to the obtention of NL-3 full sequence, shown in Figure 5.

35 Figure 6 presents a comparison of the amino acid sequence of the new NEP-like enzymes and Table I shows the extent of identity between members of the family.

Cellular distribution of NL-1, NL-2 and NL-3 peptidases

Determining the tissue distribution of NL-1, NL-2 and NL-3 may provide clues to identify the peptidergic systems in which they are involved. It will be particularly

interesting to compare the tissue distribution of these peptidases with that of NEP and the ECEs to determine whether or not the physiological functions of NL-1 and/or NL-2 and/or NL-3 may overlap those of NEP and/or ECEs.

5 *In situ* hybridization (ISH), using our mouse cDNA, was used to determine the spatial and temporal expression of NL-1 during mouse development, as done previously for PEX (Ruchon et al., 1998)). Serial sections of whole foetal (12, 15 and 19 dpc) and adult mice (1, 3 and 6 days old) were hybridized with an [³⁵S]-labeled RNA probe. Figure 7 shows a section of mouse testis which was the only tissue identified to express NL-1 by this technique. Cells of seminiferous tubules are specifically labeled
10 but spermatids located near the center of the tubule showed strongest labeling. These cells are in the last stage of maturation into spermatozooids. The presence of NL-1 in testis has now been confirmed by Northern analysis of mouse tissues (see Fig. 10). Other tissues express NL-1, when analyzed by RT-PCT, which is a more sensitive assay (not shown).

15 A similar approach was used to determine the localization of NL-3 using the mouse EST obtained from ATCC. Figure 8 shows sections of whole mouse at 17 days of embryonic development and 4 days post-natal. Several tissues are expressing this putative peptidase including brain, where it is associated with neurons (Figure 9), spinal chord, liver, spleen and bones. Labeling was stronger in bones from *Hyp* mouse, an animal model for hypophosphatemic rickets (Figure 8). In bones, NL-3 was found
20 to be expressed by osteoblasts (not shown).

 Northern blotting experiments were performed on several tissues with NL-2 and NL-3 probes. A Human Multiple Tissues Northern Blot (Clontech) was hybridized with specific probes. A single RNA band of approximately 4.0 kb was revealed by the probe
25 for NL-2. Expression of NL-2 is restricted to brain and spinal cord (not shown). However, RT-PCR has shown the presence of this enzyme in testis (not shown).

 A single RNA band of approximately 3.0 kb was detected with the specific probe for NL-3 (not shown). NL-3 expression was observed mainly in ovary, spinal cord and adrenal gland.

30 Chromosomal localisation of the human gene for NL-2 and NL-3

 As a mean to get clues on the function of the new metallopeptidases in vertebrates, we have localized the new cDNAs on human chromosomes, in order to look for a possible link between the gene locus and mapped genetic diseases in humans. To do so, we have mapped the NL-2 and NL-3 genes by high-resolution
35 fluorescence *in situ* hybridization (FISH). NL-2 was localized to chromosome band 1p36. Consistent with the cellular distribution of NL-2 in humans, genetic diseases of the CNS such as dyslexia, neural tube defect, neuroblastoma, neuronal type of Charcot-Marie-Tooth disease have all been mapped in this region and represent potential targets for a role of NL-2 in humans. NL-3 was localized to chromosome band

2q37. Consistent with a role of NL-3 in bones, a form of Albright hereditary osteodystrophy was mapped to the same chromosomal locus (Phelan et al., 1995).

In view of the foregoing, NL-2 and NL-3 are metallopeptidases that are assumed to be immediately useful as markers for a disease or disorder associated with human chromosomal locus 1p36 and 2q37, respectively. Their localization on a chromosome band associated with known diseases suggests that they may be expressed or co-expressed with one or more genes, as a cause or a consequence of disease development. It is possible that these enzymes are up or down regulated, alone or along with other genes involved in a disease. Therefore, antibodies or other ligands specific to NL-2 or NL-3 may be used for a diagnostic purpose, as well as primers or probes in diagnostic assays using nucleic acid hybridization or amplification techniques. Otherwise, primers or probes directed against the nucleic acids of NL-2 and NL-3 would be useful to map the mutations of a gene located in close proximity and involved in the disease. Therefore, no matter which exact function NL-2 and NL-3 gene products have, their chromosomal localization provides one diagnostic utility. This localization as well as the tissular distribution provide information as to the disease and tissue to be investigated to elucidate the exact function of these enzymes.

NL-1 resembles NL-2, sharing with the latter about 80% homology in the amino sequence and sharing structural characteristics such as the furin recognition sequence located at the proximal end of the ectodomain. NL-2 might be the human homologue of mouse NL-1. If such was the case, these two proteins would have a substantial degree of divergence and, maybe, different profiles of activity varying from one species to another.

Chromosomal localization of NL-1 was determined in mouse genome by Single Strand Conformational Polymorphism (SSCP) in collaboration with The Jackson Laboratory Backcross DNA Panel Mapping Resource. NL-1 was localized to the distal region of mouse chromosome 4 which corresponds to human chromosome region 1p36 where is located NL-2 gene. This reinforces our hypothesis that NL-1 and NL-2 are species variants.

Production of antibodies against NLs

Antisera collected from injected animals were first tested by immunoblotting on GST-antigen fusion proteins produced in *E. coli*. Antiserum from one rabbit recognized the NL-1-related polypeptide and antisera from one mouse and one rabbit reacted with the NL-3-related polypeptide (results not shown). The anti NL-1 antiserum and the mouse anti NL-3 antiserum, which appeared more specific than the rabbit antiserum, were next tested by immunoblotting on extracts of proteins and culture media from cells expressing NL-1 or NL-3 (see below).

Expression of NL-1 in CHO cells

The cDNA encoding the full-length NL-1 protein was cloned in the mammalian

expression vector pcDNA3-RSV and transfected in CHO cells. Stable cell lines were established by selection with the drug G418 and tested by immunoblotting for the presence of NL-1.

5 Small amounts of NL-1 were found in the extracts of transfected CHO cells (results not shown). This intracellular species was sensitive to endo H digestion, indicating that the sugar moiety was not mature and suggesting ER localization (results not shown). The culture medium of transfected CHO cells showed the presence of soluble NL-1 (Figure 10). This extracellular species was resistant to endo H suggesting true transport through the late secretory pathway. The cDNA sequence of NL-1
10 predicts a type II transmembrane protein. The mechanism by which NL-1 is transformed into a soluble protein is not known presently. However, examination of the amino acid sequence revealed the presence of a putative furin cleavage site from residue 58 to 65 (Figure 3). A similar site is present in NL-2 sequence.

The soluble form of NL-1 was assayed for activity using [³H]-Tyr-(D)Ala₂-Leu-enkephalin and bradykinin as substrates. Figure 11 shows that NL-1 can degrade the
15 enkephalin substrate ($K_m = 18 \pm 10 \mu M$) and that this activity can be inhibited by phosphoramidon ($IC_{50} = 0.9 \pm 0.3 nM$) and thiorphan ($K_m = 47 \pm 12 nM$), a general inhibitor of enzymes of the NEP family. Bradykinin is also a substrate for NL-1 (not shown).

Use of NL-1 amino-terminal domain to promote secretion

20 The observation that NL-1 ectodomain was secreted, possibly through cleavage of the transmembrane segment by furin, raised the possibility to promote secretion of exogenous proteins that could be spliced to NL-1 amino-terminal domain (from initiator methionine to the furin site). To test this hypothesis, the ectodomain of NL-3 (from the third cysteine to the end) was spliced to NL-1 amino-terminal domain using a PCR
25 strategy and the recombinant DNA cloned in expression vector pRcCMV. The fusion protein was expressed by transfection of the vector in COS-1 and HEK 293 cells. The culture media of transfected cells was analyzed by immunoblotting using the mouse antiserum against NL-3. Figure 12 shows the presence of NL-3 in the spent culture media of both COS-1 and HEK 293 cells. This result shows that NL-1 amino-terminal
30 domain can be used to promote secretion of exogenous proteins.

The soluble form of NL-3 was assayed for activity using [³H]-Tyr-(D)Ala₂-Leu-enkephalin as substrate. No activity was found.

The previous experiment showed that it was possible to use the amino-terminal domain of NL-1 to promote secretion of an otherwise membrane attached protein
35 ectodomain. To verify whether the same strategy could be used to promote secretion of small peptides, a PCR strategy was used to splice human β -endorphin to the amino-terminal domain of NL-1 and the recombinant DNA was cloned in vector pRcCMV. The fusion protein was expressed by transfection of the vector in COS-1 and HEK 293 cells. The culture media of transfected cells was collected 48h after transfection and

the peptides purified as described previously (Noël et al., 1989). The presence of β -endorphin in the extracts was detected by radioimmunoassay. The results showed that both COS-1 and HEK 293 cells produced approximately 100 pg of β -endorphin per ml of culture medium. Therefore, the N-terminus of LN-1 and NL-2 which ends with a
5 furin-recognition site will be useful to produce the soluble form of a protein of interest.

SECTION 2)

MATERIALS AND METHODS

10 DNA manipulations

All DNA manipulations, phage library screening, and plasmid preparations were performed according to standard protocols (Ausubel 1988; Sambrook 1989). Site-directed mutagenesis was performed using a PCR-based strategy as described previously (Le Moual 1994).

15

mRNA purification and RT-PCR protocol for identification of new members of the neprilysin family

mRNAs were prepared from mouse testis using Quick Prep Micro mRNA purification kit (Pharmacia Biotech). First strand cDNA was synthesized from 1 μ g of
20 mRNA using the First-Strand cDNA synthesis kit (Pharmacia Biotech).

Two sense primers, oligonucleotide 3817 (5'-TGGATGGAT/CGA/CIGG/AIACIA/CA-3') and oligonucleotide 3719 (5'-A/GTIGTITTT/CCCIGCIGGIA/GT/AIC/TTG/CCA-3') corresponding respectively to amino acid residues 459 to 465 and 552 to 560 of NEP sequence, and one antisense
25 primer, oligonucleotide 3720 (5'-AIICCIACCIA/TC/TA/GTCIGCIG/AC/TA/GTTT/CTC-3') corresponding to amino acid residues 646 to 654 (see Fig. 1 and 2), were synthesized. PCR was performed with 5 μ l of cDNA template and 1 μ l of Taq DNA polymerase in a final volume of 100 μ l, containing 1 mM $MgCl_2$, 2 μ M of each oligonucleotide 3817 and 3720, 200 μ M of each dNTP and 5% DMSO. Cycling profiles included an initial
30 denaturation step of 5 min at 94°C, 30 cycles of 1 min at 94°C, 1 min at 40°C and 1.5 min at 72°C, and a final extension step at 72°C for 10 min. One half of the amplified DNA was fractionated on a 2% agarose gel and fragments ranging in size between 500-700 bp were purified and resuspended in a final volume of 50 μ l. A second round of PCR was done with primers 3719 and 3720, using as template either 10 μ l of the
35 first PCR reaction or 5 μ l of the purified fragments, and the new PCR products were ligated in pCR2.1 vector (Invitrogen). Several identical clones corresponded to a potential new member of the NEP family. We called this member NL1 for NEP-like 1.

Cloning of full-length NL1 cDNA

The cloned NL1 PCR fragment was used as probe to screen a mouse testis λ Uni-ZAP™XR cDNA library (Stratagene). Twelve out of a hundred positive phages were plaque purified and subcloned into pBS SK vector (Stratagene). As the longest clone analyzed presented an incomplete ORF (pBS-NL1A), 5'RACE with primers located in
5 vector (5'-TAGTGGATCCCCCGGGCTGCAG-3', sense primer) and NL1 (5'-ACCAAACCTTTCCTGTAGCTCC-3', antisense primer, nt 1303 to 1324 of NL1; was subsequently performed on the DNA of the remaining semi-purified positive clones. Amplification was performed with 1 μ l of Vent polymerase in a final volume of 100 μ l containing 50 ng of DNA, 4 mM of $MgSO_4$, 1 μ M of each oligonucleotide, 200 μ M of
10 each dNTP and 10% DMSO. Cycling parameters included an initial denaturation step of 1 min at 94°C, 25 cycles of 30 sec at 94°C, 30 sec at 60°C and 1 min at 72°C, and an incubation of 10 min at 72°C. A PCR fragment of the expected length was subcloned into pCR2.1 vector (clone pCR-NL1A), but sequencing revealed no initiator ATG codon. A nested 5'RACE was then performed on mouse testis cDNA using the
15 Marathon Ready cDNA kit (Clontech) with sense oligonucleotides AP1 and AP2 (from the kit) and NL1 antisense oligonucleotides 5'-CCTGAGGGCTCGTTTTACAACCGTCCT-3' (nt 503 to 529 of NL1) and 5'-CTCATCCCAGGAGAAGTGTAGCAGGCT-3' (nt 475 to 502 of NL1) as recommended by the supplier. The resulting fragment was cloned into pCR2.1 vector (pCR-NL1B).
20 Since only ten bp were missing for the initiator ATG codon, we reconstructed the 5' end of the cDNA by PCR-amplifying clone pCR-NL1A with sense primer 5'-CCACCATGGTGGAGAGAGCAGGCTGGTGTGCGGAAGAAG-3' (nt 332 to 364 of NL1; the 10 missing nucleotides are underlined) and antisense primer 5'-ACCAAACCTTTCCTGTAGCTCC-3' (nt 1303 to 1324 of NL1) using Vent polymerase
25 as described above. The DNA fragment was then inserted into pCR2.1 (clone pCR-NL1C). The entire ORF was reconstituted following digestion of pBS-NL1A and pCR-NL1C with *EcoRI* and *PstI*. The 5' end of NL1 cDNA was excised from pCR-NL1C and ligated into pBS-NL1A at the corresponding sites, resulting in plasmid pBS-NL1B.

For expression studies, a *Bam*HI/*Ap*al fragment generated out of pBS-NL1B,
30 corresponding to the full length cDNA of NL1, was inserted into pCDNA3/RSV [18] vector.

Production of polyclonal antibodies

A plasmid for the production in *Escherichia coli* of a GST fusion protein with NL1
35 was constructed using pGEX-4T-3 expression vector (Pharmacia Biotechnologies). A 255 bp fragment from NL1 was amplified by PCR with Vent polymerase using sense primer 5'-GCTACGGGATCCCGTGGCCACTATGCTTAGGAA-3' (nt 1139 to 1158) and antisense primer 5'-CGATTGCTCGAGTGGGAACAGCTCGACTTCCA-3' (nt 1377 to 1396). Both pGEX-4T-3 and the PCR product were digested with *Bam*HI and *Xho*I

and ligated. The recombinant protein was produced and purified according to the supplier's instructions. Five weeks old female balb/c mice were immunized at monthly intervals for 3 months with 20 µg of the recombinant NL1 fragment in Freund's adjuvant and antisera were subsequently collected.

5

Cell culture and transfection

Human Embryonic Kidney (HEK 293) cells were maintained in DMEM medium containing 10% fetal bovine serum (FBS), and supplemented with penicillin at 60 µg/ml, streptomycin at 100 µg/ml and fungizone at 0,25 µg/ml. Transfections of cells with appropriate plasmids were performed by the calcium/phosphate-DNA co-precipitation method (Chang 1987). To establish permanent cell lines, G418 selection was initiated 48 h after the transfections at 400 µg/ml for 12 days and gradually decreased at 100 µg/ml.

LLC-PK₁ cells transfected with pRcCMV-sNEP were maintained as described previously (Lanctôt 1995).

Immunoblot analysis

For immunoblot analysis, cells were incubated for 16 h in synthetic DMEM medium containing 2mM sodium butyrate. Cellular proteins were solubilized as previously described (Dion 1995). Secreted proteins recovered in culture media were concentrated approximately 10 fold by ultrafiltration. Immunoblot analysis were performed using the NEN Renaissance kit with the polyclonal antibody specific to NL1 or the α1-antitrypsin inhibitor antibody (Calbiochem; LaJolla, CA) followed by the appropriate horseradish peroxidase-conjugated IgG (Vector Laboratories).

For the glycosylation studies, proteins were incubated with endoglycosidase H (endoH) or peptide:N-glycosidase (PNGaseF) as suggested by the distributor (NEB).

Enzymatic activity assays

NL1 activity was monitored and compared to sNEP activity using (Tyrosyl-[3,5-³H])(D-Ala₂)-Leu₅-enkephalin (50 Ci/mmol) (Research Products International Inc.), as already described (Dion 1995; Devault 1988). *K_m* values were determined by the isotope-dilution method. The inhibitory effects of phosphoramidon and thiorphan were also assessed as previously described (Dion 1995).

HPLC analysis of the hydrolysis of Leu-enkephalin

Five µg of Leu₅-enkephalin were incubated at 37°C for one hour in 50mM MES, pH 6.5, with concentrated culture medium of HEK 293 cells expressing NL1 (~300 µg of total proteins) or LLC-PK₁ cells expressing sNEP (~30 µg of total proteins), in absence or presence of 0.1 mM phosphoramidon. Hydrolysis products were separated by

reversed-phase HPLC as described previously [23]. Tyr-Gly-Gly and Phe-Leu were both identified by elution profiles of synthetic marker peptides.

Northern blot analysis

5 A mouse multiple tissue poly(A)⁺ mRNA blot (Clontech) was hybridized with a [³²P]dCTP random primer labelled probe in ExpressHyb solution (Clontech). The blot was washed according to the manufacturer's recommendations and exposed to Fuji RX film for 7 days at -80°C with intensifying screens.

10 RT-PCR screening of mouse tissues

First strand cDNA synthesis was performed with 1 µg of total RNA from mouse tissues and oligo(dT) as primer, using Gene Amp RNA PCR Core Kit (Perkin Elmer). For the PCR reactions, primers 5'-TGGCGAGAGTGTGTCAGCTATGTC-3' and 5'-CTTCCAAAATGTAGTCAGGGTAGCCAATC-3' were used with Taq polymerase. One
15 tenth of the PCR products were visualized on a 4% agarose gel.

In situ hybridization

To construct a plasmid for the synthesis of cRNA probes for ISH, pCR-NL1A was used as template to amplify a 452 bp fragment by PCR with sense primer 5'-
20 GGAGCCATAGTGACTCTGGGTGTC-3' (nt 416 to 439) and antisense primer 5'-GACGCTCAGCAGGGGCTCAGAGTC-3' (nt 842 to 865). The amplification product was inserted into pCRII vector (Invitrogen). Synthesis of riboprobes and protocols for ISH were as described previously (Ruchon 1998).

RESULTS

25

Cloning and sequence analysis of mouse NL1 cDNA

In order to isolate cDNAs for new members of the NEP family, we developed an RT-PCR strategy based on fact that NEP, ECE-1 and PHEX share regions of significant sequence identity. Following RT-PCR on testis mRNAs with nested primers,
30 a DNA fragment of approximately 300 bp was amplified. This DNA fragment was cloned and the plasmids from 24 independent colonies were sequenced: 3 clones had no insert, 4 clones had DNA fragments not related to the NEP family, 7 clones had sequences corresponding to mouse NEP and 3 clones had sequences corresponding to mouse PHEX, showing that our approach efficiently allowed the identification of
35 members of the family. Moreover, 7 identical clones had a new cDNA presenting sequence similarities to members of the NEP family. The full-length cDNA was subsequently obtained by screening a mouse testis λ cDNA library followed by 5'RACE, as described under *Materials and Methods*. Its nucleotide and deduced amino acid sequences confirm that we cloned a novel NEP-like protein, referred to

thereafter as NL1.

NL1 cDNA spans 2925 nt, including a 5'-untranslated region of 331 nt, an open reading frame of 2295 nt from nt 332 to nt 2626, and a 3'-untranslated region of 299 nt. The sequence surrounding the proposed initiator ATG conforms to the Kozak consensus (Kozak 1986). The deduced amino acid sequence of NL1 reveals a putative type II transmembrane protein of 765 amino acid residues encompassing a short N-terminal cytoplasmic tail, a unique putative transmembrane domain, and a large C-terminal extracellular domain. The ectodomain contains nine potential N-glycosylation sites (Asn-X-Ser/Thr) and ten cysteine residues corresponding to those conserved among all the members of the family, which are presumably involved in proper folding and in maintenance of the protein in an active conformation. All amino acid residues known to be part of the active site of NEP are present in NL1. The predicted protein presents greater similarities to NEP than to any other member of the family.

Although NL1 shares numerous features with proteins of the neprilysin family, a notable aspect distinguishes it from the others : the first conserved cysteine residue of the ectodomain is more distant (34 amino acid residues) from the predicted transmembrane domain in NL1 than it is in NEP (9 residues) or any other members of the family. Moreover, we noticed a putative furin cleavage site (-Arg₅₈-Thr-Val-Val-Lys-Arg₆₃-) between the end of the transmembrane domain and the first cysteine. This observation suggests that NL1 could exist as a secreted rather than a membrane-bound protein.

NL1 expression in HEK 293 cells

HEK 293 cells were transfected with pCDNA3/RSV expression vector containing NL1 cDNA, and a permanent cell line was established as described under *Materials and Methods* (HEK/NL1 cells). Immunoblotting with a polyclonal antibody showed that after 16h of culture, most NL1 was present in the culture medium with small amounts of the enzyme in the cell extract. Secreted and cell-associated NL1 had apparent molecular masses of approximately 125 and 110 kDa, respectively. To characterize the glycosylation state of NL1, we next submitted the recombinant protein to deglycosylation by peptide : N-glycosidase F (PNGase F) and endoglycosidase H (endo H). PNGase F removes high mannose as well as most complex N-linked oligosaccharides added in the Golgi complex. In contrast, endo H removes N-linked oligosaccharide side chains of the high mannose type found on proteins in the RER but which have not yet transited through the Golgi complex; thus, resistance to endo H can be used as an indication that the protein has traveled through the Golgi complex. PNGase F treatment showed that the cell-associated and secreted NL1 were N-glycosylated as their electrophoretic mobility increased following digestion. However, the secreted NL1 migrated as a doublet after PNGase F treatment, with one

band co-migrating with cell-associated form and the second having a slower rate of migration. Since untreated and endo H-digested secreted NL1 are seen as single bands by SDS-PAGE, our observation suggests that a proportion of secreted NL1 undergoes further post-RER postranslational modification that renders some of the N-linked oligosaccharides resistant to PNGase F digestion.

In contrast to secreted NL1, NL1 from cell extract was sensitive to endo H treatment. This result shows differences in the glycosylation state of the two species and suggests that the cell-associated form observed in transfected cells is an intracellular species that has not traveled through the Golgi complex.

Processing of NL1 by a subtilisin-like convertase

To determine whether a member of the mammalian subtilisin-like convertase family is responsible for NL1 presence in the culture medium of transfected cells, we co-transfected transiently HEK 293 cells with a constant amount of plasmid pCDNA3/RSV/NL1 and increasing amounts of plasmid pCDNA3/CMV/PDX (Benjannet 1997). This latter vector promotes the expression of the α 1-antitrypsin Portland variant, α 1-PDX, a known inhibitor of subtilisin-like convertases (Anderson 1993). Immunoblot analysis of the culture media of cells expressing both NL1 and α 1-PDX indicated that NL1 secretion was strongly inhibited by the presence of α 1-PDX: a relation was observed between the amounts of α 1-PDX and the level of inhibition of NL1 secretion.

To confirm that proteolysis by the subtilisin-like convertase occurred at the putative furin cleavage site identified in NL1 ectodomain (-Arg₅₈-Thr-Val-Val-Lys-Arg₆₃-), the amino acid residues Asn₆₂-Gly₆₃ were substituted for Lys₆₂-Arg₆₃ by site-directed mutagenesis in vector pCDNA3/RSV/NL1 and the mutated vector used to establish HEK 293 cells expressing the mutant protein (HEK/NL1mut cells). Immunoblot analysis of the culture media of HEK/NL1mut cells showed that the mutation totally abolished secretion of NL1. Furthermore, an additional form of NL1 with a molecular mass of 127 kDa was detected in the extract of these cells. This new species was resistant to endo H digestion and was found associated with membranes when HEK/NL1mut cells were fractionated according to Chidiac *et al.* 1996 (result not shown).

NL1 enzymatic activity

Culture media from HEK 293 and HEK/NL1 cells were tested for enzymatic activity using as substrate (Tyrosyl-[3,5-³H])(D-Ala₂)-Leu₅-enkephalin, a known NEP substrate. Activity was detected in the culture medium of HEK/NL1 cells but not in that of HEK 293 cells. This activity increased linearly with the amounts of NL1 and with the incubation period, indicating that degradation of the substrate was due to NL1 enzymatic activity.

We next characterized NL1 enzymatic parameters using the same substrate

and compared them to those of an engineered soluble form of NEP (sNEP) (Lemay 1989). NL1 affinity for D-Ala₂-Leu₅-enkephalin was slightly higher than that of sNEP as shown by their K_m values of 18 μ M and 73 μ M, respectively. Inhibition assays showed that phosphoramidon had similar effects on NL1 and sNEP activity, with IC₅₀ values of 0.9 nM and 0.5 nM respectively, and that thiorphan, a specific inhibitor of NEP, inhibited NL1 with an IC₅₀ of 47 nM, as compared with an IC₅₀ of 8 nM for NEP.

Very low levels of phosphoramidon-sensitive activity was detected in extracts of HEK/NL1 cells (data not shown) consistent with the small amounts of NL1 observed by immunoblotting.

To determine whether NL1 had cleavage site specificity similar to NEP, we incubated Leu₅-enkephalin in the presence of NL1 recovered from the medium of HEK/NL1 cells or in the presence of sNEP, and analyzed the degradation products by RP-HPLC. Peaks co-migrating with standard Tyr-Gly-Gly and Phe-Leu peptides were observed in both RP-HPLC profiles, indicating that both enzymes cleaved the substrate at the Gly₃-Phe₄ peptide bond. This enkephalin-degrading activity was completely inhibited by phosphoramidon (1 μ M).

Tissue and cellular distribution of NL1 mRNA

Tissue distribution of NL1 mRNA was determined by Northern blot analysis with a specific probe corresponding to the 5' end of the coding region of NL1 cDNA. A single transcript of 3.4 kb was detected exclusively in testis among all the mouse tissues tested. Mouse tissues were also screened by RT-PCR. Using this more sensitive technique, expression of NL1 was observed in several other tissues including heart, brain, spleen, lungs, liver and kidney. Consistent with the Northern blot results, RT-PCR analysis, although not strictly quantitative, detected more NL1 mRNA in testis than in other tissues.

To gain more insight into NL1 mRNA distribution, we examined by ISH cryostat sagittal sections from a 4-day newborn mouse, as well as sections from a 16-day old animal (p16) and adult tissues (heart, brain, spleen, lungs, liver, kidney and testis). The presence of NL1 mRNA was detected only in adult testis. Only the germinal cells in the luminal face of the seminiferous tubules were labeled. These cells were identified as round and elongated spermatids in all spermiogenesis maturational stages. Neither spermatozoa nor spermatocytes, spermatogonies or Sertoli cells were labeled. Interstitial cells were also negative. Controls were performed with sense riboprobes, which produced only nonspecific background (data not shown). The 4-day old mouse sagittal sections and all other tissues tested were negative.

DISCUSSION

The great interest in members of the Neprilysin family as putative therapeutic

targets, and the recent discovery of new members of this important family of peptidases led us to investigate whether additional members of the family remained to be identified. Using a PCR-based strategy, we cloned, from mouse testis, a partial cDNA encoding a new NEP-like enzyme that we called NL1. Analysis of the amino acid
5 sequence encoded by the full-length NL1 cDNA revealed that this member of the family resembles NEP the most: 55% identity and 74% similarity. Recently, the primary structure of a new zinc metallopeptidase from total mouse embryo was reported (Ikeda 1999). This enzyme, called SEP, is found either as a soluble or a cell-associated form due to alternative splicing. NL1 shows only 3 amino acid differences with the soluble
10 form of SEP indicating that secreted SEP and NL1 are the same enzyme. Our cloning strategy did not allow characterization of the cell-associated form of NL1 which is a minor species in mouse testis (Ikeda 1999).

The amino acid sequence of NL1 predicts a topology of a type II integral membrane glycoprotein that is similar to the other members of the family. Treatment of the
15 recombinant protein with PNGase F showed that indeed NL1 possesses N-linked carbohydrate side chains. However, it is not possible to determine precisely whether all nine putative N-glycosylation sites are used, but the 30 kDa decrease in molecular mass upon PNGase F treatment suggests that most are glycosylated. It has already been shown that all asparagine residues in a Asn-X-Ser/Thr consensus are
20 glycosylated in rabbit NEP expressed in COS-1 cells and that sugar moieties increase the stability and enzymatic activity of the protein and facilitate its intracellular transport (Lafrance 1994). Three of NEP glycosylated Asn residues (Asn 145, Asn 285 and Asn 628) are conserved in NL1 (Asn 163, Asn 303 and Asn 643). Amongst these residues, Asn 145 and Asn 628 have been reported to influence NEP enzymatic activity
25 (Lafrance 1994). In the same work, it has also been shown that the effect of sugar addition on folding and intracellular transport of NEP is a cumulative effect of all glycosylation sites rather than a contribution of any particular one. Glycosylation of NL1 may share similarities with that of NEP since we found their primary structures and enzymatic activities to be very similar.

30 Surprisingly, expression of the cDNA by transfection of HEK 293 cells showed that most of the enzyme was secreted in the culture medium. The small amount of NL1 associated with the cells was endo H-sensitive, suggesting that the cell-associated enzyme is a species that has not yet left the RER. The presence of a furin cleavage site in NL1 sequence between the predicted transmembrane domain and the first
35 conserved cysteine residue of the ectodomain led us to believe that a member of the mammalian subtilisin-like family of convertases was responsible for the presence of NL1 in the culture medium. These enzymes are involved in processing a variety of precursor proteins such as growth factors and hormones, receptors, plasma proteins, matrix metalloproteinases, metalloproteases-desintegrins and viral envelope

glycoproteins [for a review see: (Nakayama 1997). Site-directed mutagenesis of the furin cleavage site (-Arg₅₈-Thr-Val-Val-Lys-Arg₆₃-) and expression of α 1-PDX, a potent inhibitor of mammalian subtilisin-like convertases (Anderson 1993), confirmed that a member of this family of endoproteases was involved in NL1 secretion presumably by
5 cleaving in carboxy-terminus of Arg₆₃. There are only a few examples of proteins which are processed from a membrane-bound precursor to a secreted form following cleavage by subtilisin-like convertases; these include meprin and collagen XVII (Milhiet 1995; Schacke 1998). Three members of the subtilisin-like family of convertases, namely furin, PC4 and PC7, are known to be expressed in germ cells (Nakayama
10 1992; Torri 1993; Seidah 1992, 1996). Whether one of these convertases generates secreted NL1 from its membrane form is under current investigation. In any case, NL1 is the only known member of the neprilysin family that is secreted. This unique feature suggests that NL1 plays its physiological role in a context different from that of the membrane-bound peptidases, thereby diversifying the role of the peptidases of the neprilysin family. It is of interest that circulating forms of NEP in blood and urine have
15 been described, but they have generally been related to pathological or stressful conditions (Almenoff 1984; Deschodt-Lanckmann 1989; Johnson 1985; Soleilhac 1996; Aviv 1995).

We have observed in cells expressing NL1 mutated at the furin cleavage site the
20 appearance of a species resistant to digestion by endo H. This mutated protein was associated with cellular membranes. Taken together, these results indicate that NL1 is first synthesized and inserted in the RER membrane as a type II transmembrane protein. During intracellular transport, NL1 is converted to a soluble form by the action of a member of the mammalian subtilisin-like convertases. The identity of the cellular
25 compartment where this process occurs is not known. However, mammalian subtilisin-like convertases are usually active in post-Golgi compartments of the secretory pathway suggesting that processing of NL1 from the membrane bound form to the soluble form is a post-Golgi event.

Despite almost total abrogation of NL1 secretion, we observed only a slight
30 accumulation of endo H-resistant NL1 in cells either co-expressing α 1-PDX and NL1 (result not shown) or expressing mutated NL1. This observation suggests that unprocessed NL1 is rapidly degraded. A similar behavior was reported for the Notch1 receptor expressed in the furin-deficient cell line LoVo (Logeat 1998). The mechanism(s) by which these unprocessed proteins are degraded is still unknown. It
35 is interesting to point out that the spliceoform of SEP that has lost a 23 amino acid peptide, including the furin cleavage site, generates a cell-associated endo H-sensitive molecule (Ikeda 1999).

The most important observation regarding the NL1 primary structure is the conservation of residues which in NEP are essential for catalysis and binding of

substrates or inhibitors. This finding suggests that NL1 could effectively act as an endopeptidase with a catalytic mechanism similar to that of NEP. This hypothesis was supported by the demonstration that D-Ala₂-Leu₅-enkephalin, a peptide substrate often used to monitor NEP activity, was also an excellent NL1 substrate. The affinity of NL1 for D-Ala₂-Leu₅-enkephalin was even higher than that of NEP, as reflected by a K_m value 4- to 5-fold lower. Furthermore, two well known NEP inhibitors, phosphoramidon and thiorphan, also abolished NL1 activity. Phosphoramidon, which inhibits NEP as well as ECE-1 activity, albeit to a lesser extent (Turner 1996), had very similar effects on NL1 and NEP, with an IC_{50} value for NL1 varying not more than two-fold from the value determined for NEP. Thiorphan, considered to be a more specific inhibitor of NEP, also inhibited NL1 activity, with an IC_{50} six-fold greater than that for NEP. These results suggest that the active sites of NL1 and NEP are similar. This hypothesis is supported by the observation that secreted SEP degraded a set of peptides known to be NEP substrates, including substance P, bradykinin and atrial natriuretic peptide (Ikeda 1999). Taken together, these results illustrate the importance of identifying and characterizing other member of the family for the design of highly specific inhibitors.

In agreement with the enzymatic parameters demonstrating that NL1 and NEP have similar catalytic sites, we have observed that both enzymes cleaved Leu₅-enkephalin at the same peptide bond. This result suggests that NL1 hydrolyzes peptide bonds on the amino side of hydrophobic amino acid residues as does NEP (Turner 1985). However, several other peptides will have to be tested to confirm this specificity and to determine whether NL1 has dipeptidyl carboxypeptidase activity as was shown for NEP (Malfroy 1982; Bateman 1989; Beaumont 1991) and more recently for ECE-1 (Johnson 1999).

RT-PCR experiments with specific primers for the soluble and cell-associated forms of SEP showed a wide tissue distribution of the enzyme with the soluble form of SEP being predominant in testis and the cell-associated form in other tissues (Ikeda 1999). Our RT-PCR results confirmed the wide tissue distribution of NL1. However, Northern blotting and *in situ* hybridization experiments indicated that expression of NL1 is predominant in germ cells of mature testis. Interestingly, proenkephalin mRNA has been shown to be expressed in germ cells and somatic cells of the testis (Torii 1993, Seidah 1992; Kew 1989; Mehta 1994; Kilpatrick 1986, 1987). Specific functions for testicular enkephalin peptides have not yet been defined, but it is believed that they could act as intratesticular paracrine/autocrine factors. In addition to their putative role as mediators of testicular cell communication, it has also been demonstrated that proenkephalin products synthesized by spermatogenic cells during spermatogenesis are stored in the acrosome of human, hamster, rat and sheep spermatozoa and are release from sperm following acrosomal reaction (Kew 1990). It has thus been proposed that proenkephalin products may act as sperm acrosomal factors during the

fertilization process as well as intratesticular regulators secreted by spermatogenic cells. Since Leu₅-enkephalin was found to be a good substrate for NL1, opioid peptides originating from proenkephalin could serve as physiological substrate for this new enzyme. In this way, NL1 would serve to regulate the activity of these bioactive peptides.

Testis is the only tissue where the soluble form of SEP is predominant (Ikeda, 1999), suggesting a testis-specific alternative splicing. Expression of testis-specific molecular species of peptidases or prohormones, arising through diverse mechanisms, has been documented in the past (Howard 1990; Jeannotte 1987). However, the physiological significance of these testis-specific species is not always clear. In the case of NL1 or SEP, it might allow local constitutive secretion by germinal cells of an otherwise cell-associated enzyme, to regulate spermatogenesis much like several other proteolytic enzymes of the seminiferous tubules (Monsees 1998). Alternatively, it might allow accumulation in acrosome with proenkephalin peptides and release upon acrosomal reaction. More exhaustive studies concerning NL1 localization and physiological substrate identification will be needed to understand its role in the testis and possibly in the fertilization process.

Cloning of other members of the family

To find other members of the NEP-like family, we will use the same RT-PCR strategy to amplify mRNA isolated from tissues known to be regulated by peptidergic systems (brain, thymus, kidney, heart, lung, ovary, pancreas, bone, bone marrow and lymphoid cells). In fact, many of these tissues are known to express at least one member of the family and/or to control a peptidergic pathway on which peptidase inhibitors have major effects. Amplified fragments will be cloned and the resulting clones will be sequenced and compared to the sequence of known peptidases, as described above. Pairs of degenerate primers in other highly conserved regions will also be designed to increase the possibility of cloning other relevant peptidases.

DISCUSSION

As discussed above, peptidases of the NEP family known to date have often been found to play important physiological roles. This is certainly true for NEP itself, ECEs and PEX, (see review above). For this reason, some of these enzymes (as it was the case for NEP and ECE in the past) might be interesting targets for the design of inhibitors that in turn could be used as therapeutic agents in various pathological conditions. However, it is of some concern that inhibitors designed for one enzyme may also inhibit to some extent other members of the family. This lack of specificity for an inhibitor used as a therapeutic agent in the long term treatments such as those used as antihypertensive agents for instance, may cause unforeseen problems due to unwanted side effects. The objectives of the present work was to develop a strategy

to clone new members of the NEP family of peptidases. The results presented in this report clearly show that our strategy can be successful. We have determined the complete or partial nucleotide sequence of three cDNAs encoding putative enzymes of the NEP family.

5 These cDNA sequences are valuable tools and may be used to:

Produce antibodies

As shown in the present work, knowledge of NL cDNA sequences can be used to raise specific antibodies. For example but not exclusively, regions of less homology between the peptidases (amino acid residues 50 to 450) can be used to synthesize peptides whose sequences are deduced from the translation of the cDNAs, and/or bacterially-expressed fragments of the cDNAs fused for example but not exclusively to GST may be purified and injected into rabbits or mice for polyclonal or monoclonal antibody production. These antibodies can be used to:

- 15 - identify by immunohistochemistry the peptidergic pathways in which the peptidases are functioning;
- study the physiopathology of NL-enzymes by immunoblotting or immunohistochemistry on samples of biological fluids or biopsies;
- set up high through put screening assays to identify NL-enzymes inhibitors. This can be done for example but not exclusively by using the antibodies to attach the NL-enzymes to a solid support;
- 20 - purify NL-enzymes with said antibodies by immunoprecipitation or affinity chromatography by identifying antibodies capable of selectively binding to the NL-enzymes in one set of conditions and releasing it in another set of conditions typically involving a large pH or salt concentration change without denaturing the NL-enzyme;
- 25 - identify antibodies that block NL-enzymes activities and use them as therapeutic agents. Blocking antibodies can be identified by adding antisera or ascite fluid to an *in vitro* enzymatic assay and looking for inhibition of NL-enzymes activities. Blocking antibodies could then be injected to normal or disease model animals to test for *in vivo* effects.

Derive specific RNA or DNA probes

As shown in the present work, knowledge of the nucleotide sequence of the members of the NEP-family allows nucleotide sequence comparisons and facilitate the design of specific RNA or DNA probes by methods such as but not exclusively molecular cloning, *in vitro* transcription, PCR or DNA synthesis. The probes thus obtained can be used to:

- 35 - derive specific probes or oligonucleotides for RNA and DNA analysis, such as Northern blot and *in situ* hybridization, chromosome mapping

- or PCR testing. These probes could be used for genetic testing of normal or pathological samples of biological fluids or biopsies;
- make vectors for gene knock-out or knock-in in mice. The long range PCR technique and/or screening of a mouse genomic library with probes derived from the 5'-end of the cDNAs can be used to isolate large exon/intron regions. We will then substitute one or more of the cloned genomic DNA exons for the neomycin resistance expression cassette for producing homologous recombination and knock-out mice. Alternatively, cDNAs coding for NLs will be used to overexpressed each of these enzymes in transgenic mice. The cDNAs will be cloned downstream from a promoter sequence, and injected in fertilised mouse eggs. Depending on specific questions to be answered, the chosen promoter sequence will allow expression of the peptidases either in every tissues or in a cell- or tissue-specific manner. Injected eggs will be transferred into foster mothers and the resulting mice analysed for peptidase expression;
 - replace defective NL genes in a gene therapy strategy. The NL full length cDNAs could be cloned under the control of a constitutive or inducible promoter having a narrow or wide range of tissue expression and introduced with appropriate vectors in subjects having defective genes;
 - synthesise oligonucleotides that could be used to interfere with the expression of the NLs. For example but not exclusively, oligonucleotides with antisens or ribozyme activity could be developed. These oligonucleotides could be introduced in subjects as described above;
 - isolate other members of the family. Screening cDNA and/or genomic libraries with these cDNA probes at low stringency may allow to clone new members of the NEP-like family. Alternatively, alignment of the sequences may allow one to design specific degenerate oligonucleotide primers for RT-PCR screening with mRNA from tissues such as but not exclusively, the hearth and the brain.

Production of recombinant NL-enzymes

As shown in the present work, recombinant active NL-enzymes can be obtained by expression of NL-cDNAs in mammalian cells. From past experience with neprilysin, another member of the family (Devault *et al.*, 1988; Fossiez *et al.*, 1992; Ellefsen *et al.*, submitted), expression can also be performed in other expression systems after cloning of NL-cDNAs in appropriate expression vectors. These expression systems may include but not exclusively the baculovirus/insect cells or larvae system and the *Pichia*

pastoris-based yeast system. Production of recombinant NL-enzymes includes the production of naturally occurring membrane bound or soluble forms of the proteins or genetically engineered soluble forms of the enzymes. The latter can be obtained by substituting the cytosolic and trans-membrane domain by a cleavable signal peptide such as that of proopiomelanocortin, but not exclusively, as done previously (Lemay *et al.*, 1989) or by transforming by genetic manipulations the non-cleavable signal peptide membrane anchor domain into a cleavable signal peptide, as done previously (Lemire *et al.*, 1997) or by fusion of the ectodomain of NL-enzymes to the amino-terminal domain (from the initiator methionine to amino acid residue 300) of naturally occurring soluble NLs such as, but not exclusively, NL-1 as done in this work.

These recombinant NLs could be used to:

- find a substrate. A substrate can be identified using one of the following.
- Screening of existing bioactive peptides. Peptides are incubated in the presence of NL-enzymes and subsequently analysed by HPLC for degradation. Degradation is observed by disappearance of the peak of substrate and the appearance of peaks of products;
- Screening phage libraries specifically designed for the purpose (phage display library). Each phage expresses at its surface, as part of its coat protein, a random peptide sequence preceded by a peptide sequence recognisable by an antibody or any other sequence-recognizing agent. This latter sequence serves to attach the phage to a solid support. Upon addition of the NL-enzyme the random sequences that are NL substrate are cleaved, releasing the phage. After several rounds of cleavage, the phage sequence is determined to identify the peptide segment recognized by the enzyme.
- Extract of the tissue where the enzyme is expressed is collected and prepared for chromatographic analysis (HPLC, capillary electrophoresis or any other high resolution separation system) by denaturing the extracted proteins with a solvent (acetonitrile or methanol). The extract is subjected to chromatographic separation. The same extract is incubated with the enzyme for a period sufficient to observe a difference between the 2 chromatograms. The regions with the identified changes are collected and subjected to mass spectrometric analysis to determine the peptide compositions.
- Small peptide libraries are prepared with a fluorophore at one extremity and a quencher group at the other (Meldal *et al* Methods in molecular biology 1998,87). The substrate can be identified using a strategy described in Apletalina *et al* (JBC (1998)273, 41, 26589-95). For each

- 5 hexapeptide library, the identity of one residue at one position remains constant while the rest is randomized (for a total of $6 \times 20 = 120$ individual libraries). Each library is made-up of 3.2 million different members and is identified both by the position of the constant residue along the hexapeptide, and its identity. The NL-enzyme is added to each library and the fluorescence is recorded. The data is organized to identify the libraries producing the most fluorescence for each position along the hexapeptide. This arrangement suggests the identity of important residues at each position along the hexapeptide. Hexapeptide representing the best suggestions are prepared and tested in a similar fashion. From this set, the hexapeptide with the best fluorescence is selected.
- 10
- 15 - set up enzymatic assays. An enzymatic assay consists in the addition of the above-identified substrate to the enzyme in constant conditions of pH, salts, temperature and time. The resulting solution is assayed for the hydrolysed peptide or for the intact peptide. This assay can be realized with specific antibodies, HPLC or, when self-quenched fluorescence tagged peptides are used (Meldal et al), by the appearance of fluorescence. The enzyme may be in solution or attached to a solid substrate;
- 20
- 25 - identify inhibitors. Inhibitors can be identified from synthetic libraries, biota extracts and from rationally designed inhibitors using X-ray crystallography and substituent activity relationships. Each molecule or extract fraction is tested for inhibitory activity using the enzymatic test described above. The molecule responsible for the largest inhibition is further tested to determine its pharmacological and toxicological properties following known procedures. The inhibitor with the best distribution, pharmacological action combined with low toxicity will be selected for drug manufacturing. Pharmaceutically acceptable formulation of the inhibitor or its acceptable salt will be prepared by mixing with known excipients to produce tablets, capsules or injectable solutions. Between 1 and 500mg of the drug is administered to the patients;
- 30
- 35 - inject the native or soluble purified NL-enzymes into subjects. In the case of disease or pathologies caused by a lack or decrease in NL activity, the purified NL could be injected intravenously or otherwise in patients. Alternatively, immobilized NL-enzymes could be introduced at the site of orthopedic surgery or implantation of devices in bones or dental tissues.

Secretion of foreign proteins and peptides

As shown in the present work, the amino-terminal domain of NL-1 (from the initiator methionine to the furin site) can be used to promote the secretion of a foreign protein (in this case the ectodomain of NL-3 and β -endorphin).

The amino-terminal domain of NL-1 but also of other naturally occurring soluble NL-enzymes could be used to:

- promote production and secretion of foreign proteins. This can be achieved by genetically fusing sequences coding for said foreign proteins downstream from and in phase with the amino-terminal of NL-1. These chimeric constructs could be introduced with the help of appropriate vectors in any of the expression systems mentioned above for protein production and secretion;
- promote production and secretion of bioactive peptides. Sequences encoding small bioactive peptides such as but not exclusively β -endorphin, the enkephalins, substance P, atrial natriuretic peptide (ANF) and osteostatine, could be fused immediately downstream and in phase the furin site of NL-1. These DNA constructs could be used as described above to produce bioactive peptides.
- serve as model to design artificial (non-naturally occurring) proteins or protein segments (protein vectors) to promote secretion of proteins or peptides. These protein vectors can be constructed to resemble a secreted protein. In this case they would be assembled of an endoplasmic reticulum signal peptide, a spacer of varying length and a furin cleavage site to which the protein or peptide destined for secretion can be fused. The total length of the spacer, furin cleavage site and protein or peptide destined for secretion must be at least 70 amino acid residues. Alternatively, such protein vectors could be assembled to resemble a type II membrane protein. In this case they would comprise from the amino to the carboxy-terminus a cytosolic domain of varying length, a transmembrane domain that also acts as a signal peptide, an extracellular segment of varying length and a furin cleavage site to which the protein or peptide destined for secretion can be fused. The total length of the extracellular segment, furin cleavage site and protein or peptide destined for secretion must be at least 70 amino acid residues.

Therapeutic applications of NL-enzymes

The inappropriate processing of endogenous peptides causes several diseases. The inappropriate processing may result from pathologic

concentration of the enzyme itself, its substrate or other elements of the biochemical machinery downstream from the controlling enzyme. In this context it is possible to help the patient by managing the activity of the controlling enzyme.

- 5 - NL-enzymes have been localized to the brain and may be involved in the improper processing of β -amyloid precursor. Inhibitions of this process by drugs prepared as above, will help patients with Alzheimer disease as well as other patient suffering from diseases caused by plaque formation;
- 10 - NL-enzymes may be involved in the improper processing of other peptides involved in neurological diseases, pain or psychiatric disorders. Appropriately designed inhibitors will help in the management of such diseases;
- 15 - NL-1 is found in testis and is associated with spermatozoid maturation. Peptides improperly processed by the enzyme may lead to infertility. The addition of NL-1 ex-vivo to seminal liquid or immature spermatozooids taken directly from testis during an in-vitro fertilization procedure will increase fertility. Conversely, the use of a small-molecule inhibitor or removal of NL-1 with an antibody could increase fertility during an in-vitro fertilization procedure. The administration of a NL-1 inhibitor may increase or decrease the fertility potential. This inhibitor is formulated and administered as described above.
- 20 - NL-3 is found in ovaries and may be involved in the processing of a peptide involved in the maturation of eggs. The addition of NL-3 ex-vivo to immature eggs taken directly from ovaries during an in-vitro fertilization procedure will increase fertility. Conversely, the use of a small-molecule inhibitor or removal of NL-3 with an antibody could increase fertility during an in-vitro fertilization procedure. This inhibitor is formulated and administered as described above;
- 25 - NL-3 is found in bones. The improper processing of peptides by the enzyme may result in bone disease or abnormal phosphate metabolism. Administration of an inhibitor, as described above, will allow the disease management.
- 30 - NL-3 is found in bones. The improper processing of peptides by the enzyme may result in bone disease or abnormal phosphate metabolism. Administration of an inhibitor, as described above, will allow the disease management.

35

TABLE I

Extend of amino acid sequence identity between members of the NEP-like family

	hNEP	hPEX	hECE-1A	hECE-2	hKELL	sNL-1	hNL-2	hNL-3
hNEP	100*							

5

hPEX	35	100						
hECE-1A	39	38	100					
hECE-2	36	37	62	100				
hKELL	23	24	30	31	100			
sNL-1	55	39	39	39	26	100		
hNL-2	54	39	39	39	26	77	100	
hNL-3	35	32	37	37	28	36	34	100

*: percentage of sequence identity

REFERENCE LIST

- Almenoff, J., Teirstein, A.S., Thornton, J.C., and Orlowski, M. (1984) *J. Lab. Clin. Med.* **103**, 420-431
- Anderson, A.D., Thomas, L., Hayflick, J.S., and Thomas, G. (1993) *J. Biol. Chem.* **268**, 24887-24891
- 5 Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., and Smith, J.A. (1988). *Current protocols in molecular biology* (New York: Wiley Interscience).
- Aviv, R., Gurbanov, K., Hoffman, A., Blumberg, S., and Winaver, J. (1995) *Kidney Int.* **47**, 855-860
- 10 Bateman, R.C., Jackson D., Slaughter, C.A., Unnithan, S., Chai, Y.G., Moomaw, C., and Hersh, L.B. (1989) *J. Biol. Chem.* **264**, 6151-6157
- Bawab, W., Aloyz, R.S., Crine, P., Roques, B.P., DesGroseillers, L. (1993) *Biochem. J.* **296**, 459-465
- Beaumont, A., Le Moual, H., Boileau, G., Crine, P., and Roques, B.P. (1991). Evidence that both arginine 102 and arginine 747 are involved in substrate binding to neutral endopeptidase (EC 3.4.24.11). *J.Biol.Chem.* **266**, 214-220.
- 15 Benjannet, S., Savaria, D., Laslop, A., Munzer, J.S., Chrétien, M., Marcinkiewicz, M., and Seidah, N.G. (1997) *J. Biol. Chem.* **272**, 26210-26218.
- Chen, C. and Okayama, H. (1987) *Mol. Cell Biol.* **7**, 2745-2752
- 20 Chidiac, P., Nouet, S., and Bouvier, M. (1996) *Mol. Pharmacol.* **50**, 662-669
- Corvol, P. and Williams, T.A. (1997). Biochemical and molecular aspects of angiotensin I-converting enzyme. In *Cell-surface peptidases in health and disease*. A.J. Kenny and C.M. Boustead, eds. (Oxford, UK: BIOS Scientific Publishers Ltd), pp. 99-117.
- 25 Crine, P., Dion, N., and Boileau, G. (1997). Endopeptidase-24.11. In *Cell-Surface Peptidases in Health and Disease*. A.J. Kenny and C.M. Boustead, eds. (Oxford: BIOS Scientific Publishers), pp. 79-98.
- Crine, P., LeGrimellec, C., Lemieux, E., Labonté, L., Fortin, S., Blachier, A. and Aubry, M. (1985) The production and characterization of a monoclonal antibody specific for the 94000 dalton enkephalin-degrading peptidase from rabbit kidney brush border. *Biochem. Biophys Res. Commun.* **131**, 255-261.
- 30 Deschodt-Lanckman, M., Michaux, F., de Prez, E., Abramowicz, D., Vanherweghem, J.L., and Goldman, M. (1989) *Life Sci.* **45**, 133-141
- Devault, A., Nault, C., Zollinger, M., Fournié-Zaluski, M.-C., Roques, B.P., Crine, P., and Boileau, G. (1988a). Expression of neutral endopeptidase (enkephalinase) in heterologous COS-1 cells. Characterization of the recombinant enzyme and evidence for a glutamic acid residue at the active site. *J.Biol.Chem.* **263**, 4033-4040.
- 35 Devault, A., Sales, V., Nault, C., Beaumont, A., Roques, B.P., Crine, P., and Boileau,

- G. (1988b). Exploration of the catalytic site of endopeptidase 24.11 by site-directed mutagenesis. Histidine residues 583 and 587 are essential for catalysis. *FEBS Lett.* **231**, 54-58.
- 5 Dion, N., Le Moual, H., Crine, P., and Boileau, G. (1993). Kinetic evidence that His-711 of neutral endopeptidase 24.11 is involved in stabilization of the transition state. *FEBS Lett.* **318**, 301-304.
- Dion, N., Le Moual, H., Fournié-Zaluski, M.C., Roques, B.P., Crine, P., and Boileau, G. (1995). Evidence that Asn⁵⁴² of neprilysin (EC 3.4.24.11) is involved in binding of the P₂' residue of substrates and inhibitors. *Biochem.J.* **311**, 623-627.
- 10 Holmes, M.A. and Matthews, B.W. (1982). Structure of thermolysin refined at 1.6 Å resolution. *J.Mol.Biol.* **160**, 623-639.
- Hooper, N.M. (1994). Families of zinc metalloproteases. *FEBS Lett.* **354**, 1-6.
- Howard, T.E., Shai, S.Y., Langford, K.G., Martin, B.M., Bernstein, K.E. (1990) *Mol. Cell. Biol.* **10**, 4294-4302
- 15 Ikeda, K., Emoto, N., Raharjo, S.B., Nurhantari, Y., Saiki, K., Yokoyama, M., Matsuo, M. (1999) *J. Biol. Chem.* **274**, 32469-32477
- Jeannotte, L., Burbach, J.P., Drouin, J. (1987) *Mol. Endocrinol.* **1**, 749-757
- Johnson, A.R., Coalson, J.J., Ashton, J., Larumbide, M., and Erdos, E.G. (1985) *Am. Rev. Respir. Dis.* **132**, 1262-1267
- 20 Johnson, G.D., Stevenson, T., and Ahn, K. (1999) *J. Biol. Chem.* **274**, 4053-4058
- Endocrinol.* **3**, 1191-1196
- Kenny, J. (1993). Endopeptidase-24.11: Putative substrates and possible roles. *Biochem.Soc.Trans.* **21**, 663-668.
- Kew, D., Jin, D.F., Kim, F., Laddis, T., and Kilpatrick, D.L. (1989) *Mol. J. Biol Chem.* **266**, 214-220
- 25 Kew, D., Muffly, K.E., and Kilpatrick, D.L. (1990) *Proc. Natl. Acad. Sci. U.S.A* **87**, 9143-9147
- Kilpatrick, D.L., Borland, K., and Jin, D.F. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5695-5699
- 30 Kilpatrick, D.L. and Millette, C.F. (1986) *Proc. Natl. Acad. Sci. U.S.A* **83**, 5015-5018
- Kozak, M. (1986) *Cell* **44**, 283-292
- Lafrance, M.H., Vezina, C., Wang, Q., Boileau, G., Crine, P., and Lemay, G. (1994) *Biochem. J.* **302**, 451-454.
- Lancôt, C., Fournier, H., Howell, S., Boileau, G., and Crine, P. (1995) *Biochem. J.* **305**, 165-171
- 35 Le Moual, H., Devault, A., Roques, B.P., Crine, P., and Boileau, G. (1991). Identification of glutamic acid 646 as a zinc-coordinating residue in endopeptidase-24.11. *J.Biol.Chem.* **266**, 15670-15674.
- Le Moual, H., Dion, N., Roques, B.P., Crine, P., and Boileau, G. (1994). Asp650 is

- crucial for catalytic activity of neutral endopeptidase 24-11. *Eur.J.Biochem.* **221**, 475-480.
- Lemay, G., Waksman, G., Roques, B.P., Crine, P., and Boileau, G. (1989). Fusion of a cleavable signal peptide to the ectodomain of neutral endopeptidase (EC 3.4.24.11) results in the secretion of an active enzyme in COS-1 cells. *J.Biol.Chem.* **264**, 15620-15623.
- Logeat, F., Bessia, C., Brou, C., LeBail, O., Jarriault, S., Seidah, N.G., and Israël, A. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 8108-8112.
- Malfroy, B., and Schwartz, J.C. (1982) *Biochem. Biophys. Res.Com.* **106**, 276-285
- 10 Marie-Claire, C., Ruffet, E., Antonczak, S., Beaumont, A., O'Donohue, M., Roques, B.P., and Fournié-Zaluski, M.C. (1997). Evidence by site-directed mutagenesis that arginine 203 of thermolysin and arginine 717 of neprilysin (neutral endopeptidase) play equivalent critical roles in substrate hydrolysis and inhibitor binding. *Biochemistry* **36**, 13938-13945.
- 15 Mehta, N.D., Don, J., Zinn, S.A., Millette, C., Wolgemuth, D.J., and Kilpatrick, D.L. (1994) *Endocrinology* **135**, 1543-1550
- Milhiet, P.E., Chevallier, S., Corbeil, D., Seidah, N.G., Crine, P., and Boileau, G. (1995) *Biochem. J.* **309**, 683-688
- Monsees, T.K., Gornig, M., Schill, W.B., and Miska, W. (1998) *Andrologia* **30**, 185-191
- 20 Nakayama, K. (1997) *Biochem. J.* **327**, 625-635
- Nakayama, K., Kim, W.S., Torii, S., Hosaka, M., Nakagawa, T., Ikemizu, J., Baba, T., and Murakami, K. (1992) *J. Biol. Chem.* **267**, 5897-5900
- Phelan, M.C., Rogers, R.C., Clarkson, K.B., Bowyer, F.P., Levine, M.A., Estabrooks, L.L., Severson, M.C., and Dobyns, W.B. (1995). Albright hereditary osteodystrophy and del(2) (q37.3) in four unrelated individuals. *Am.J.Med.Genet.* **58**, 1-7.
- 25 Raut, R., Rouleau, J.L., Blais, C. Jr, Gosselin, H., Molinaro, G., Sirois, M.G., Lepage, Y., Crine, P., and Adam, A. (1999) Bradykinin metabolism in the postinfarcted rat heart: role of ACE and neutral endopeptidase 24.11. *Am. J Physiol.* **276**, H1769-H1779.
- 30 Ruchon, A.F., Marcinkiewicz, M., Siegfried, G., Tenenhouse, H.S., DesGroseillers, L., Crine, P., and Boileau, G. (1998). Pex mRNA is localized in developing mouse osteoblasts and odontoblasts. *J.Histochem.Cytochem.* **46**, 459-468.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular cloning, a laboratory manual* (New-York: Cold Spring Harbor Laboratory Press).
- 35 Schacke, H., Schumann, H., Hammami-Hausli, N., Raghunath, M., and Bruckner-Tuderman, L. (1998) *J. Biol. Chem.* **273**, 25937-25943
- Seidah, N.G. and Chrétien, M. (1995). Pro-protein convertases of subtilisin/kexin family. *Methods Enzymol.* **244**, 175-188.

- Seidah, N.G., Day, R., Hamelin, J., Gaspar, A., Collard, M.W., and Chretien, M. (1992) *Mol. Endocrinol.* **6**, 1559-1570
- Seidah, N.G., Hamelin, J., Mamarbachi, M., Dong, W., Tadros, H., Mbikai, M., Chretien, M., and Day, R. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 3388-3393
- 5 Soleilhac, J.M., Lafuma, C., Porcher, J.M., Auburtin, G., and Roques, B.P. (1996) *Eur. J. Clin. Invest.* **26**, 1011-1017
- Torii, S., Yamagishi, T., Murakami, K., and Nakayama, K. (1993) *FEBS Lett.* **316**, 12-16
- 10 Turner, A.J. (1997a). Endothelin-converting enzymes. In *Cell-surface peptidases in health and disease*. A.J. Kenny and C.M. Boustead, eds. (Oxford, UK: BIOS Scientific Publishers Ltd.), pp. 137-153.
- Turner, A.J. and Tanzawa, K. (1997b). Mammalian membrane metallopeptidases: NEP, ECE, KELL, and PEX. *FASEB J.* **11**, 355-364.
- Turner, A.J., and Murphy, L. (1996) *Biochem. Pharmacol.* **51**, 91-102
- 15 Turner, A.J., Matsas, R., and Kenny, A.J. (1985) *Biochem. Pharmacol.* **34**, 1347-1356
- Vijayaraghavan, J., Kim, Y.-A., Jackson, D., Orlowski, M., and Hersh, L.B. (1990). Use of site-directed mutagenesis to identify valine-573 in the S'1 binding site of rat neutral endopeptidase 24.11 (enkephalinase). *Biochemistry* **29**, 8052-8056.

What is claimed is:

1. A neutral endopeptidase-like metallopeptidase which is found upon probing tissue nucleic acids with degenerate oligonucleotides derived from a conserved sequence located on either side of a sequence His-Glu-Xaa-Xaa-His, wherein Xaa is any amino acid, which has a sequence selected from the amino acid sequences shown in Figures 3, 4, and 5, a fragment thereof, and a variant thereof sharing at least about 80% homology with said sequence.
2. A metallopeptidase as defined in claim 1, which has the amino acid sequence shown in Figure 3.
3. A metallopeptidase as defined in claim 1, which has the amino acid sequence shown in Figure 4.
4. A metallopeptidase as defined in claim 1, which has the amino acid sequence shown in Figure 5.
5. A nucleic acid encoding the metallopeptidase of any one of claims 1 to 4.
6. A recombinant vector comprising the nucleic acid defined in claim 5.
7. A recombinant host cell expressing the nucleic acid of claim 5.
8. A method for producing a metallopeptidase as defined in any one of claims 1 to 4, which comprises the steps of culturing a recombinant host cell as defined in claim 7 in a growth supportive medium and recovering said metallopeptidase from said host cell or the culture medium.
9. A method for screening new molecules related to neutral endopeptidase (NEP), which comprises the steps of:
 - aligning nucleotidic sequences of NEP and of known molecules related to NEP;
 - assessing consensus sequences on either side of a sequence comprising His-Glu-Xaa-Xaa-His sequence, wherein Xaa is any amino acid;
 - synthesising degenerate sequences of said consensus sequences;
 - contacting said degenerate sequences with the nucleic acids of a panel of samples susceptible to express said new molecules, in conditions such that a hybridization complex can form between the nucleic acids of samples and the degenerate sequences;

- detecting the formation of said hybridization complex as an indication of a sample which comprises a molecule related to NEP; and
 - sequencing the nucleic acid comprising said hybridization complex;
- whereby a new sequence sharing homology with NEP is a new molecule related to NEP.

10. A method as defined in claim 9, wherein said degenerate sequences are selected from Figure 2.

11. An oligonucleotide selected from those in Figure 2.

12. A composition of matter comprising one or more of the oligonucleotides of claim 11.

13. A recombinant vector comprising a nucleic acid encoding the N-terminal part of the amino acid sequence shown in Figure 3 or 4, which N-terminal part terminates with a furin-recognition sequence.

14. A host cell transformed with the recombinant vector of claim 13.

15. A method for producing a soluble form of a membrane protein of interest having a C-terminal ectodomain, said soluble form essentially consisting of said ectodomain, which comprises:

- obtaining nucleic acids encoding essentially the ectodomain;
- fusing the nucleic acids in phase with the C-terminal end of the N-terminal part of the recombinant vector defined in claim 13;
- having the fused nucleic acids to be expressed in a host cell in the presence of a culture medium, which host cell expresses or is made to express furin; and
- recovering said soluble form in the culture medium.

16. A method as defined in claim 15, wherein said protein of interest is NL-3 or β -endorphin.

17. The soluble form of the metallopeptidase defined in any one of claims 1 to 4 which soluble form essentially consists of the ectodomain of said metallopeptidase.

18. A composition comprising the soluble metallopeptidase of claim 17.

19. An oligonucleotide derived from the nucleic acid defined in claim 4, which oligonucleotide has at least 12 nucleic acids in length.

20. An antibody directed against the metallopeptidase defined in any one of claims
5 1 to 4.

21. A method for detecting the presence or amount of a metallopeptidase as defined in any one of claims 1 to 4 in a sample, which comprises the steps of contacting said sample with the antibody defined in claim 20, in conditions such that
10 an immune complex is formed between said antibody and said metallopeptidase, and detecting the presence or amount of an immune complex as an indication of the presence or amount of said metallopeptidase in said sample.

22. A method of detecting the presence or amount of a metallopeptidase as defined
15 any one of claims 1 to 4 in a sample, which comprises the steps of: contacting said sample susceptible to comprise a target nucleic acid with the nucleic acid defined in claim 5 or with an oligonucleotide as defined in claim 19 in conditions such that a hybridization complex can form between the target nucleic acid of the sample and said nucleic acid or oligonucleotide encoding said metallopeptidase, and detecting the
20 formation of such a hybridization complex as an indication of the presence of said metallopeptidase in said sample.

1/22

	1	10	20	30	
NEP1-HU	MGK.....	SESQMDITDINT..	PKPKKKQRWTPLEI.....	SLSVLVLLL	
	*	*	*	*	
PEX-HUM	MEA.....	ETG....	SSVET..	GKKANRGTRIALVV.....	FVGGTLVLG
	**	*	*	*	*
KELL-HU	MEGGDQSEEEPRERSQAGGMGTLSQESTPEERLPVEGSRPWAV....	ARRVLTAILIL.			
	*	*	*	*	*
ECE1-HU	MSTYKRATLDEEDLVDSLSEGDAYPNGLQVNFHSPRSGQRCWAARTQVEKRLVVLVLLA				
consens	M		T	P	L

	40	50	60	70	80	90
NEP1-HU	TIIAVTMIALYA.TYDD...	GICKSSDCIKSAARLIQNMDATTEPCTDFFKYACGGWLKR				
	**	*	*	*	*	*
PEX-HUM	TILFLVSQGLLSLQAKQ...	EYCLKPECIEAAAAILSKVNLSVDPCDNFFRFACDGWISN				
	*	*	*	*	*	*
KELL-HU	.GLLLCFSVLLFYNFQNCGRPCEPETSVCCLDLRDHYLASGNTSVAPCTDFFSFACG...RA					
	**	*	*	*	*	*
ECE1-HU	AGLVACLAALGI.QYQTRSPSVCLSEACVSVTSSILSSMDPTVDPCHDFFSYACGGWIKA					
consens	L	L	C	C	L	V PC DFF ACGGW

	100	110	120	130	140	150
NEP1-HU	NVIPETSSRYGNFDILRDELEVVLKDVLPQEP..	KTEDIVAVQKAKALYRSCINESAIDSR				
	*	*	*	*	*	*
PEX-HUM	NPIPEDMPSYGVYPWLRHNVDLKLKELLEKSISRRRDTEAIQKAKILYSSCMNEKAIEKA					
	*	*	*	*	*	*
KELL-HU	KETNNS.....	FQELATKNKNRLRRILEVQ..	NSWHPGSGEKAFQFYNSCMDTLAIEAA			
	*	*	*	*	*	*
ECE1-HU	NPVPDGHRSRWGTFNSNLWEHNQAI IKHLLENS.TA.SVSEAERKAQVYYRACMNETRIEEL					
consens	N	P	G F L	LK LE	A KA	Y SCMNE AIE

	160	170	180	190	200
NEP1-HU	GGEPLLKLLPDI.YGWP..	VATENWEQKYGAS.WTAEKAIQALNSKYGKKVLINLFVGTD			
	***	*	*	*	*
PEX-HUM	DAKPLLHILRHSPFRWPVLESNIGPEGVWSEKFSLLQTLATFRGQYSNSVFIRLYVSPD				
	**	*	*	*	*
KELL-HU	GTGPLRQVIEEL.....	GGWRISGKWTSLNFN..	RTLRLMSQYGHFPFFRAYLGPH		
	**	*	*	*	*
ECE1-HU	RAKPLMELIERL.....	GGWNITGPWAKDNFQ..	DTLQVVTAHYRTSPFFSVVVSAD		
consens	PL		G W	F	TL Y F YV D

	220	230	240	250	260
NEP1-HU	DKNSVNHVIHIDQPRGLPSR.DYYECTGIYKEACTAYVDFMISVARLIRQEERLPI.DE				
	**	*	*	*	*
PEX-HUM	DKASNEHILKLDQATLSLAVREDYLDNSTEAKSYRDALYKFMVDTAVL.....	LGA.NS			
	**	*	*	*	*
KELL-HU	PASPHTPVIQIDQPEFDVPLKQDQEQKI.YAQIFRE.YLTYLNQLGTL.....	LGG.DP			
	***	*	*	*	*
ECE1-HU	SKSNSNSNVIQVDQSGGLGPSRDYYLNKTENEKVLTG.YLNYMVQLGKL.....	LGGGDE			
consens	K S	VI DQ	L LP R DY	K Y M	L LG D

FILE 1

2/22

270 280 290 300 310 320
 NEP1-HU NQLALEMNKVMLEKEIANATAKPEDRNDPMLLYNKMTLAQIQNNFSLEINGKPFWSLNF
 * * * * *
 PEX-HUM SRAEHDMKSVLRLEIKIAEIMI PHENRTSEAMY.NKMNISELSAMIP.....QFDWLGY
 * * * * *
 KELL-HU SKVQEHSSLSISITSRLFQFLRPLEQORRAQGKLFQMVITIDQLKEMAP.....AIDWLSC
 * * * * *
 ECE1-HU EAIRPQMQQILDFETALANITIPQEKRRDEELIYHKVTAAELQTLAP.....AINWLFF
 * * * * *
 consens M E A P E R K T L P WL

330 340 350 360 370 380
 NEP1-HU TNEIMSTVNISITNEEDVVVYA....PEYLTCLKPIILTKYSARDLQNLMSWRIFIMDLVSS
 * * * * *
 PEX-HUM IKKVIDTRLYPHLKDISPSENVVVRVPQYFKDLFRILGSEKKTIANYLVRMVMYSRIPN
 * * * * *
 KELL-HU LQATFTPMSLSPSQSLVVHDVEYL...KNMSQLVEEMLLKQRFELQSHMILGLVVTLSPA
 * * * * *
 ECE1-HU LNTIFYPVEINESEPIVVYDKEYL...EQISTLINT...TDRCLLNNYMIWNLVKRTSSF
 * * * * *
 consens V L L N M W V

390 400 410 420 430
 NEP1-HU LSRTYKESR...NAFRKALYGTT.SETATWRRCANVYVNGNMENAVGRLYVEAAFAGESK
 *** * * * *
 PEX-HUM LSRRFQYRW...LEFSRVIQGT.TLLPQWDKCVNFIESALPYVVGKMFVDVYFQEDKK
 * * * * *
 KELL-HU LDSQFQEARRLSOKLRELTEQPPMPARPRWMKCVEETGTFFEPTLAALFVREAFGPSTR
 * * * * *
 ECE1-HU LDQRFQDA....DEKFMEVMYGTCKTCLPRWKFCVSDTENNLGFALGPMFVKATFAEDSK
 * * * * *
 consens L FQ F GT P W CV G FV F K

440 450 460 470 480 490
 NEP1-HU HVVEDLIAQIREVFIQTLD.DLTWMDAETKKRAEEKALAIKERIGYPDDIVSNDKNLNE
 * * * * *
 PEX-HUM EMMEELVEGVRWAFIDMLEKENEWMDAGTKRKAKEKARAVLAKVGYPE.FIMNDTHVNED
 * * * * *
 KELL-HU SAAMKLFTAIRDALITRLR.NLPWMNEETONMAQDKVAQLQVEMGASE.WALKPELARQE
 * * * * *
 ECE1-HU SIATEIILEIKKAFESLS.TLKWMDEETRKSAKEKADAIYNMIGYPN.FIMDPKELDKV
 * * * * *
 consens L IR AFI L L WMD ET A EKA A GYP
 (4)
 (1A/B)

500 510 520 530 540 550
 NEP1-HU YLELNYKEDEYFENIIQNLKFSQSKQLKKLREKVDKDEWISGAAVVNAFYSSGRNQIVFP
 * * * * *
 PEX-HUM LKAIKFSEADYFGNVLTQRKYLAQSDFFWLRLKAVPKTEWFTNPTTVNAFYSSASTNQIRFP
 * * * * *
 KELL-HU YND.IQLGSSFLQSVLSCVRSRLRIVQSFLQPHQHRWKVSPWDVNAYYSVSDHVVVFP
 * * * * *
 ECE1-HU FNDYTAVPDLYFENAMRFFNFWSRVTDQLRKAPNRDQWSMTPPMVNAYYSPTKNEIVFP
 * * * * *
 consens YF N LR W P VNA YS N IVFP
 (2A/B)

1 (cont'd)

3/22

	560	570	580	590	600	610
NEP1-HU	AGILQPPFFSAQQ.SNSLNYGGIGMVIGHEITHGFDDNGRNFNKDGLVDWWTQOSASNF					
	** ** *					** ** *
PEX-HUM	AGELQKPPFFWGTEYPRSLSYGAIGVIVGHEFTHGFDNNGRKYDKNGNLDPPWWSTESEEF					
	** ** *					** ** *
KELL-HU	AGLLQPPFFHPGY.PRAVNFGAAGSIMAHELLHIFYQL...LLPGGCL....ACDNHAL					
	** ** *					** ** *
ECE1-HU	AGILQAPFYTRSS.PKALNFGGIGVVVGHELTHAFDDQGREYDKDGNLRPWWKNSSVEAF					
consens	AG LQ PFF	P	LN G IG	GHE TH FD GR	K G L WW	S F

	620	630	640	650	660	670
NEP1-HU	KEQSQCMVYQYGNFSWDLAGGQHLNGINTLGENIADNGGGLQAYRAYQNYI..KKNQ.EE					
	** ** *					** ** *
PEX-HUM	KEKTKCMINQYSNYYWK.KAGLNVKGKRTLGENIADNGGGLREAFRAYRKWINDRRQGLEE					
	* * *					* * *
KELL-HU	QEAHLCLKRHYAAF..PLPSRTSFNDSLTFLENAADVGGGLAIALQAYSKRL..LRHH.GE					
	* * *					* * *
ECE1-HU	KRQTECMVEQYSNY..SVNG.EPVNGRHTLGENIADNGGGLKAAYRAYQNWV..KKNQ.AE					
consens	KE CM QY N		NG TL	GENIADNGGGL	A RAY	G E

	680	690	700	710	720	730
NEP1-HU	KLLPGLDLNHHKQLFFLNFAQVWCGTYRPEYAVNSIKTDVHSPGNFRIIGTLQNSAEFSEA					
	*****					*****
PEX-HUM	PLLPGITFTNNQLFFLSYAHVRCNSYRPEAAREQVQIGAHSPQFRVNGAISNFEEFQKA					
	**					**
KELL-HU	TVLPSLDLSPQQIFFRSYAQVMCRKPSQDSH.....DTHSPPHLRVHGPLSSTPAFARY					
	*****					*****
ECE1-HU	HSLPTLGLTNNQLFFLGFAQVWCSVRTPESSHEGLITDPHSPSRFRVIGLSLSNSKEFSEH					
consens	LP L L QLFFL	AQV C	PE	D HSP	FRV G LSN	EF

	740	750
NEP1-HU	FHCRKNSYNNPEKK.CRVW	
	* * * * *	
PEX-HUM	FNCPPNSTMNRGMDSCRLW	
	* * *	
KELL-HU	FRCARGALLNPSSR.CQLW	
	*** * *	
ECE1-HU	FRCPPGSPMNPCHK.CEVW	
consens	F C S MNP	C W

Fig 1 (cont'd)

4/22

PRIMER	SEQUENCE
(1A)	5'-TGGATGGAT/CGA/CIGG/AIACIA/CA-3'
(1B)	5'-TGGATGGAT/CGA/CIGG/AIACIA/CG-3'
(2A)	5'-A/GTIGTITTT/CCCIGCIGGIA/GT/AIC/TTA/TCA-3'
(2B)	5'-A/GTIGTITTT/CCCIGCIGGIA/GT/AIC/TTG/CCA-3'
(3)	5'-AIIICCICCIA/TC/TA/GTCIGCIG/AC/TA/GTTT/CTC-3'
(4)	5'-GAT/CAAT/CT/CTIGAT/CGAA/GT/CTIAAT/CTGGATGG-3'
(5)	5'-T/CT/CACCAIATICT/GA/GCATCG/TT/CTTCATIGGG/ATG-3'

FILE - 2

Sequence of NL-1 cDNA from mouse

[illegible]

met val glu arg ala gly trp cys arg lys lys ser pro gly phe val glu tyr gly leu met val leu leu leu leu leu gly ala
1 ATG GTG GAG AGA GCA GGC TGG TGT CGG AAG AAG TCC CCA GGC TTC GTG GAG TAT GGG CTG ATG GTG CTG CTG TTG CTG GGA GCC

ile val thr leu gly val phe tyr ser ile gly lys gln leu pro leu leu thr ser leu leu his phe ser trp asp glu arg thr val
91 ATA GTG ACT CTG GGT GTC TTC TAC AGC ATA GGG AAG CAG CTG CCC CTC TTA ACT AGC CTG CTA CAC TTC TCC TGG GAT GAG AGG ACG GTT

val lys arg ala leu arg asp ser ser leu lys ser asp ile cys thr thr pro ser cys val ile ala ala arg ile leu glu asn
181 GTA AAA CGA GCC CTC AGG GAT TCA TCA CTG AAA AGT GAC ATC TGC ACC ACC CCA AGC TGT GTG ATA GCA GCT GCC AGA ATC CTC GAA AAC

met asp gln ser arg asn pro cys glu asn phe tyr gln tyz ala cys gly gly trp leu arg his val ile pro glu thr asn ser
271 ATG GAC CAA TCG AGG AAC CCC TGT GAA AAC TTC TAC CAG TAC GCC TGC GGA GGC TGG CTG AGG CAC CAC GTG ATC CCA GAG ACC AAC TCC

arg tyr ser val phe asp ile leu arg asp glu leu glu val ile leu lys gly val leu glu asp ser thr ser gln his arg pro ala
361 CGA TAC AGC GTC TTT GAC ATC CTC CGG GAC GAG CTG GAG GTT ATC CTC AAA GGG GTG CTG GAG GAT TCC ACT TCC CAG CAT CGC CGG GCC

val glu lys ala lys thr leu tyr arg ser cys met asn gln ser val ile glu lys arg asp ser glu pro leu leu ser val leu lys
451 GTG GAG AAG GCC AAG ACA CTA TAT CGC TCC TGC ATG ATG AAC CAA AGT GTG ATC GAG AAG AGA GAC TCT GAG CCC CTG CTG AGC GTC TTA AAA

met val gly gly trp pro val ala met asp lys lys trp asn glu thr met gly leu lys trp glu leu glu arg gln leu ala val leu asn
541 ATG GTA GGA GGT TGG CCT GTG GCC ATG GAT AAG TGG AAC GAG ACC ATG GGC CTC AAG TGG GAA CTG GAG CGA CAG TTG GCT GTG TTG AAC

ser gln phe asn arg arg val leu ile asp leu phe ile trp asn asp gln asn ser ser arg his val ile tyr ile asp gln pro
631 TCG CAG TTC AAC AGG CGG GTC CTC ATC GAC CTC TTC ATC TGG AAT GAC GAC CAG AAC TCC AGC CGG CAT GTC ATC TAC ATA GAC CAG CCC

thr leu gly met pro ser arg glu tyr tyr phe gln glu asp asn his lys val arg lys ala tyr leu glu phe met thr ser val
721 ACC TTG GGC ATG CCA TCC CGG GAG TAC TAT TTC CAG GAG AAC AAC CAC AAC GTA CGG AAA GCC TAC CTG GAG TTC ATG ACG TCA GTG

5/22

4-11

6/22

ala thr met leu arg lys asp gln asn leu ser lys glu ser ala met val arg glu glu met ala glu val leu glu leu glu thr his
 811 GCC ACT ATG CTT AGG AAA GAC CAG AAC CTG TCC AAG GAG AGC GCC ATG GTG CGG GAG ATG GCG GAG GTG GAA CTG GAG ACG CAT
 leu ala asn ala thr val pro gln glu lys arg his asp val thr ala leu tyr his arg met asp leu met glu leu gln glu arg phe
 901 CTG GCC AAC GCC ACA GTC CCC CAG CAG AAG AAA AGG CAT GAT GTC ACT GCC CTG TAC CAC CGA ATG GAC CTG ATG GAG CTA CAG GAA AGG TTT
 gly leu lys gly phe asn trp thr leu phe ile gln asn val leu ser ser val glu val glu phe pro asp glu glu val val val
 991 GGT CTG AAG GGG TTT AAC TGG ACT CTC TTC ATA CAA AAC GTC GTG TCT TCT TCT GAA GTC GAG CTG TTC CCA GAT GAG GTG GTG GTC
 tyr gly ile pro tyr leu glu asn leu glu asp ile ile asp ser tyr ser ala arg thr met gln asn tyr leu val trp arg leu val
 1081 TAC GGC ATC CCC TAC CTG GAG AAT CTG GAG GAT ATC ATT GAT AGC TAC TCA GCA CGG ACC ATG CAG AAC TAC CTG GTA TGG CGC CTG GTG
 leu asp arg ile gly ser leu ser gln arg phe lys glu ala arg val asp tyr arg lys ala leu tyr gly thr thr val glu glu val
 1171 CTA GAT CGA ATT GGC AGC CTG AGC CAG AGA TTC AAA GAG GCG CGT GTG GAC TAC CGC AAG GCG CTG TAC GGC ACC GTG GAG GAG GTA
 arg trp arg glu cys val ser tyr val asn ser asn met glu ser ala val gly ser leu tyr ile lys arg ala phe ser lys asp ser
 1261 CGC TGG CGA GAG TGT GTC AGC TAT GTC AAC AGT AAC ATG GAG AGC GCC GTG GGC TCC CTC TAC ATC AAG CGG GCC TTC TCC AAG GAC AGC
 lys ser thr val arg glu leu ile glu lys ile arg ser val phe val asp asn leu asp glu leu asn trp met asp glu glu ser lys
 1351 AAG AGC ACG GTC AGA GAG CTG ATT GAG AAG ATA AGG TCC GTG TTT GTG GAT AAC CTG GAT GAG CTG AAC TGG ATG GAC GAA TCC AAG
 lys lys ala gln glu lys ala met asn ile arg glu gln ile gly tyr pro asp tyr ile leu glu asp asn asn lys his leu asp glu
 1441 AAG AAG GCC CAG GAA AAG GCC ATG AAT ATA CGG GAA CAG ATT GGC TAC CCT GAC TAC ATT TTG GAA GAT AAC AAT AAA CAC CTG GAT GAG
 glu tyr ser ser leu thr phe tyr glu asp leu tyr phe glu asn gly leu gln asn leu lys asn asn ala gln arg ser leu lys lys
 1531 GAA TAC TCC AGT TTG ACT TTC TAT GAG GAC CTG TAT TTT GAG AAC GGA CTT CAG AAC CTC AAG AAC AAT GCC CAG AGG AGC CTC AAG AAG
 leu arg glu lys val asp gln asn leu trp ile ile gly ala ala val val asn ala phe tyr ser pro asn arg asn gln ile val phe
 1621 CTT CGG GAA AAG GTG GAC CAG AAT CTC TGG ATC ATC GGG GCT GCA GTG GTC AAT GCA TTC TAC TCC CCA AAC AGA AAC CAG ATC GTC TTT


 (cont'd)

7/22

```

pro ala gly ile leu gln pro pro phe phe ser lys asp gln pro gln ser leu asn phe gly gly ile gly met val ile gly his glu
1711 CCA GCA GGG ATT CTC CAG CCG CCC TTC TTC AGC AAG GAC CAA CCA CAG TCC TTG AAT TTT GGG GGC ATC GGG ATG GTG ATT GGG CAC GAG
ile thr his gly phe asp asp asn gly arg asn phe asp lys asn gly asn met leu asp trp trp ser asn phe ser ala arg his phe
1801 ATC ACA CAC GGC TTT GAT GAT AAT GGT CGT AAC TTT GAC AAG AAC GGC AAC ATG CTG GAC TGG TGG AGT AAC TTC TCG GCC CGG CAC TTC
gln gln gln ser gln cys met ile tyr gln tyr gln gln tyr gln phe ser trp gln leu ala asp asn gln asn val asn gly phe ser thr leu
1891 CAA CAG CAG TCG CAA TGC ATG ATC TAT CAG TAC GGC AAC TTC TCT TGG GAA CTA GCA GAC AAC CAG AAT GTG AAC GGA TTC AGT ACC CTC
gly glu asn ile ala asp asn gly gly val arg gln ala tyr lys ala tyr leu arg trp leu ala asp gly gly lys asp gln arg leu
1981 GGG GAG AAC ATT GCC GAC AAC GGA GGT GTG CGA CAG GCA TAC AAG GCT TAC CCG TGG CTG GCT GAT GGC GGC AAA GAT CAG CGA CTG
pro gly leu asn leu thr tyr ala gln leu phe phe ile asn tyr ala gln val trp cys gly ser tyr arg pro glu phe ala val gln
2071 CCG GGA CTG AAC CTG ACC TAT GCC CAG CTT TTC ATC AAC TAT GCC CAG GTG TGG TGT GGG TCC TAT AGG CCG GAG TTC GCC GTC CAG
ser ile lys thr asp val his ser pro leu lys tyr arg val leu gly ser leu gln asn leu pro gly phe ser glu ala phe his cys
2161 TCC ATC AAG ACG GAC GTC CAC AGT CCT CTT AAG TAC AGG GTG CTG GGC TCA CTA CAG AAC CTA CCA GGC TTC TCT GAG GCA TTC CAC TGC
pro arg gly ser pro met his pro met lys arg cys arg ile trp ***
2251 CCA CGA GGC AGC CCC ATG CAC CCC ATG AAG CGA TGT CGC ATC TGG TAG CCAAGGCTGAGCTATGCTGGGCCACGCCGCCACCCAGAGGCTTCGGCGATG
2354 GTGTAGCTGGCAGAGATGTGCAGGTCTTTGCTGAAGGCCGAGCCAGCCAGCCCTCCGCCGCTAGAGTGATAGCCACCCGCCACACCCGGGATGAGTGTGCGCGTC
2473 CTGCGGCCCTCAGGCCAGTGAGGGTCAGCAGCCCGAGGAGCAGTCAGCTGCCTTCCACCCCTCTCCATAGTGTGTGGCTAAATGTCTCGAGCTTCAGACTTGAGCTAAGTAAACGC
2925 TTC

```

- 3 (cont'd)

Sequence of NL-2 cDNA from humans

0 met val glu ser ala gly arg ala gly gln lys arg pro gly phe leu glu gly gly leu leu leu leu leu leu val thr
 1 GTG GGG ATG GTG GAG AGC GCC GGC CGT GCA GGG CAG AAG CCG CCG GGG TTC CTG GAG GGG GGG CTG CTG CTG CTG CTG GTG ACC
 58 ala ala leu val ala leu gly val leu tyr ala asp arg arg gly lys gln pro arg leu ala ser arg leu cys phe leu gln glu
 91 GCT GCC CTG GTG GCC TTG GGT GTC CTC TAC GCC GAC CCG AGA GGG AAG CAG CTG CCA CGC CTT GCT AGC CGG CTG TGC TTC TTA CAG GAG
 88 glu arg thr phe val lys arg lys pro arg gly ile pro glu ala gln glu val ser glu val cys thr thr pro gly cys val ile ala
 181 GAG AGG ACC TTT GTA AAA CGA AAA CCC CGA GGG ATC CCA GAG GCC CAA GAG GTG AGC GAG GTC TGC ACC ACC CCT GGC TGC GTG ATA GCA
 118 ala ala arg ile leu gln asn met asp pro thr thr glu pro cys asp phe tyr gln phe ala cys gly gly trp leu arg arg his
 271 GCC GCC AGG ATC CTC CAG AAC ATG GAC CCG ACC ACG GAA CCG TGT GAC GAC TTC TAC CAG TTT GCA TGC GGA GGC TGG CTG CGG CAC
 148 val ile pro glu thr asn ser arg tyr ser ile phe asp val leu arg asp glu leu glu val ile leu lys ala val leu glu asn ser
 361 GTG ATC CCT GAG ACC AAC TCA AGA TAC AGC ATC TTT GAC GTC CTC CGC GAC GAG CTG GAG GTC ATC CTC AAA GCG GTG CTG GAG AAT TCG
 178 thr ala lys asp arg pro ala val glu lys ala arg thr leu tyr arg ser cys met asn gln ser val ile glu lys arg gly ser gln
 451 ACT GCC AAG GAC CGG CGT GTG GAG AAG GCC AGG ACG CTG TAC CGC TCC TGC ATG AAC CAG AGT GTG ATA GAG AAG CGA GGC TCT CAG
 208 pro leu leu asp ile leu glu val val gly gly trp pro val ala met asp arg trp asn glu thr val gly leu glu trp glu leu glu
 541 CCC CTG CTG GAC ATC TTG GAG GTG GGA GGC TGG CCG GTG GCG ATG GAC AGG TGG AAC GAG ACC GTA GGA CTC GAG TGG GAG CTG GAG
 238 arg gln leu ala leu met asn ser gln phe phe asn arg arg val leu ile asp leu phe ile trp asn asp asp gln asn ser ser arg his
 631 CGG CAG CTG GCG CTG ATG AAC TCA CAG TTC AAC AGG CGC GTC CTC ATC TGG AAC GAC GAC CAG AAC TCC AGC CGG CAC
 268 ile ile tyr ile asp gln pro thr leu gly met pro ser arg glu tyr phe asn gly gly ser asn arg lys val arg glu ala tyr
 721 ATC ATC TAC ATA GAC CAG CCC ACC TTG GGC ATG CCC TCC CGA GAG TAC TAC TTC AAC GGC GGC AGC AAC CCG AAG GTG CGG GAA GCC TAC
 298 leu gln phe met val ser val ala thr leu leu arg glu asp ala asn leu pro arg asp ser cys leu val gln glu asp met val gln
 811 CTG CAG TTC ATG GTG TCA GTG GCC ACG TTG CTG CGG GAG GAT GCA AAC CTG CCC AGG GAC AGC TGC CTG GTG CAG GAG GAC ATG GTG CAG
 328

8/22



9/22

val leu glu leu thr gln thr gln ala lys ala thr val pro gln glu glu arg his asp val ile ala leu tyr his arg met gly leu
 901 GTT CTG GAG CTG GAG ACA CAG CAG CTG GCC AAG GCC ACG GTA CCC CAG GAG GAG AGA CAC GAG CTG ATC GCC TTG TAC CAC CGG ATG GGA CTG 358

glu glu leu gln ser gln phe gly leu lys gly phe asn trp thr leu phe ile gln thr val leu ser ser val lys ile lys leu leu
 991 GAG GAG CTG CAA AGC CAG TTT GGC CTG AAG GGA TTT AAC TGG ACT CTG TTC ATA CAA ACT GTG CTA TCC TCT GTC AAA ATC AAG CTG CTG 388

pro asp glu glu val val tyr gly ile pro tyr leu gln asn leu glu asn ile ile asp thr tyr ser ala arg thr ile gln asn
 1081 CCA GAT GAG GAA GTG GTG GTC TAT GGC ATC CCC TAC CTG CAG AAC CTT GAA AAC ATC ATC GAC ACC TAC TCA GCC AGG ACC ATA CAG AAC 418

tyr leu val trp arg leu val leu asp arg ile gly ser leu ser gln arg phe lys asp thr arg val asn tyr arg lys ala leu phe
 1171 TAC CTG GTC TGG CGC CTG GTG GAC CGC ATT GGT AGC CTA AGC CAG AGA TTC AAG GAC ACA CGA GTG AAC TAC CGC AAG GCG CTG TTT 448

gly thr met val glu glu val arg trp arg glu cys val gly tyr val asn ser asn met glu asn ala val gly ser leu tyr val arg
 1261 GGC ACA ATG GTG GAG GAG GTG CGC TGG CGT GAA TGT GTG GGC TAC GTC AAC AGC AAC ATG GAG AAC GCC GTG GGC TCC CTC TAC GTC AGG 478

glu ala phe pro gly asp ser lys ser met val arg glu leu ile asp lys val arg thr val phe val glu thr leu asp glu leu gly
 1351 GAG GCG TTC CCT GGA GAC AGC AAG AGC ATG GTC AGA GAA CTC ATT GAC AAG GTG CGG ACA GTG TTT GTG GAG ACG CTG GAC GAG CTG GGC 508

trp met asp glu glu ser lys lys ala gln glu lys ala met ser ile arg glu gln ile gly his pro asp tyr ile leu glu glu
 1441 TGG ATG GAC GAG TCC AAG AAG AAG CCG CAG GAG AAG GCC ATG AGC ATC CGG GAG CAG ATC GCG CAC CCT GAC TAC ATC CTG GAG GAG 538

met asn arg arg leu asp glu glu tyr ser asn leu asn phe ser glu asp leu tyr phe glu asn ser leu gln asn leu lys val gly
 1531 ATG AAC AGG CGC CTG GAC GAG TAC TCC AAT CTG AAC TTC TCA CAG GAC CTG TAC TTT GAG AAC AGT CTG CAG AAC CTC AAG GTG GGC 568

ala gln arg ser leu arg lys leu arg glu lys val asp pro asn leu trp ile ile gly ala ala val val asn ala phe tyr ser pro
 1621 CCC CAG CCG AGC CTC AGG AAG CTT CCG GAA AAG GTG GAC CCA AAT CTC TGG ATC ATC GCG GCG CTG GTC AAT CCG TTC TAC TCC CCA 598

asn arg asn gln ile val phe pro ala gly ile leu gln pro phe phe ser lys glu gln pro gln ala leu asn phe gly gly ile
 1711 AAC CGA AAC CAG ATT GTA TTC CCT GCG GCG ATC CTC CAG CCC TTC TTC AGC AAG GAG CAG CCA CAG GCC TTG AAC TTT GGA GGC ATT 628

FIG - 4 (cont'd)

10/22

gly met val ile gly his glu ile thr his gly phe asp asp asp lys asn gly asn met met asp trp trp ser
 1801 GCG ATG GTG ATC GGG CAC GAG ATC ACG CAC GGC TTT GAC GAC AAT GGC CCG AAC TTC GAC AAT GCG AAC ATG ATG GAT TGG TGG AGT
 658
 asn phe ser thr gln his phe arg glu ser glu cys met ile tyr gln tyr gly asn tyr ser trp asp leu ala asp glu gln asn
 1891 AAC TTC TCC ACC CAG CAC TTC CCG GAG CAG TCA CAG TGC ATG ATC TAC CAG TAC GGC AAC TAC TCC TGG CAC CTG GCA GAC GAA CAG AAC
 688
 val asn gly phe asn thr leu gly glu asn ile ala asp asp asp gly val arg gln ala tyr lys ala tyr leu lys trp met ala glu
 1981 CTG AAC GCA TTC AAC ACC CTT GGG GAA AAC AAT GCT GAC AAC GCA GGC GTG CCG CAA GCC TAT AAG GCC TAC CTC AAG TGG ATG GCA GAG
 718
 gly gly lys asp gln gln leu pro gly leu asp thr his glu gln leu phe phe ile asn tyr ala gln val trp cys gly ser tyr
 2071 GGT GGC AAG GAC CAG CAG CTG CCC GGC CTG GAT CTC ACC CAT GAG CAG CTC TTC TTC ATC AAC TAT GCC CAG GTG TGG TGC GGG TCC TAC
 748
 arg pro glu phe ala ile gln ser ile lys thr asp val his ser pro leu lys tyr arg val leu gly ser leu gln asn leu ala ala
 2161 CGG CCC GAG TTC GCC ATC CAA TCC ATC AAG ACA GAC GTC CAC AGT CCC CTG AAG TAC AGG GTA CTG GGG TCG CTG CAG AAC CTG GCC GCC
 770
 phe ala asp thr phe his cys ala arg gly thr pro met his pro lys glu arg cys arg val trp ter
 2251 TTC GCA GAC ACG TTC CAC TGT GCC CGG ACC ACC CCC ATG CAC CCC AAG GAG CGA TGC CGC GTG TGG TAG CCA AGG CCC TGC CGC GCT GTG
 2341 CGG CCC ACG CCC ACC CGC TGC TCG GAG GCA TCT GTG CGA AGG TGC AGC TAG CGG CGA CCC AGT GTA CGT CCC GCC CCG AAC CAT GCC
 2431 AAG CCT GCC TGC CAG GCC TCT GCG CCT GGC CTA GGG TGC AGC CAC CTG CCT GAC ACC CAG GGA TGA GCA GTG TCC AGT GCA GTA CCT GGA
 2521 CCG GAG CCC CCT TCA CAG ACA CCC GCG GGG CTC AGT GCC CCC GTC ACA ACT CTG TAG AGA CAA TCA ACT GTG TCC TGC CCA CCC TTC AAG
 2611 GTG CAT TGT CTT CCA GTA TCT ACA GCT TCA GAA CTT GAG CTA AGT AAA TGC TTT CAA AGA AAA AAA

~~FIG - 4~~ (cont'd)

Sequence of NL-3 cDNA from human

1 GGC GCT GGG AGA CAC CGG ACG CCC GGT CGG CTG TGC GGC TCA GGC CCC CGC TCG GCC CGA CCC GCT CGG TCA CCG CGG GCT CGG GCG
 91 CGC ACC TGC CGG CTG CGG CCC CAG GGC CAT GCG-GAG GCC CAC GAG GAG GCC GGC CAC GCG CAT CCC GTA GCC CAG GTG GCC CAG GTC
 181 TGC ACC GCG GCG GCC TCG GCG CCG ATG GAG CCC CCG TAT TCG CTG ACG GCG CAC TAC GAT GAG TTC CAA GAG GTC RAG TAC GTG AGC AGC
 met glu pro pro tyr ser leu thr ala his tyr asp glu phe gln glu val lys tyr val ser arg
 271 TGC GGC GCG GGG GGC GCG GCG TCC CTC CCC CCG GGC TTC CCG TTG GGC GCT GCG CGC AGC GCC ACC GGG GCC TCC GGG CTG
 cys gly ala gly gly ala arg gly ala ser leu pro pro gly phe pro leu gly ala ala arg ser ala thr gly ala arg ser gly leu
 361 CGC CGC TGG AAC CGG CGC GAG GTG TGC CTG CTG TCG GGG CTG GTG TTC GCC GGC CAC gly leu cys ala ile leu ala ala met leu ala leu
 pro arg trp asn arg arg glu val cys leu leu ser gly leu val phe ala ala gly leu cys ala ile leu ala ala met leu ala leu
 451 AAG TAC CTG GGC CCG CTC GCG GCC GGC GGC GGC TGT CCC GAG GGC TGC CCT GAG CGC AAG GCC TTC GCG CGC GCC GCT CGC TTC CTG
 lys tyr leu gly pro val ala ala gly gly ala cys pro glu gly cys pro glu arg lys ala phe ala arg ala ala arg phe leu
 ala ala asn leu asp ala ser ile asp pro cys gln asp phe tyr ser phe ala cys gly gly tip leu arg arg his ala ile pro asp
 541 GCC GCC AAC CTG GAC GCC AGC ATC GAC CCA TGC CAG GAC TTC TAC TCG TTC GCC TGC GGC GGT TGG CTG CGG CGC CAC GCC ATC CCC
 asp lys leu thr tyr gly thr ile ala ala ile gly glu gln asn glu glu arg leu arg arg leu ala arg pro gly gly pro
 631 GAC AAG CTC ACC TAT GGC ACC ATC GCG GCA ATC GGC GAG CAA AAC GAG GAG CGC CTA CCG CGC CTG CGG CGC GGT GGT GCG CCT
 gly gly ala ala gln arg lys val arg ala phe phe arg ser cys leu asp met arg glu ile glu arg leu gly pro arg pro met leu
 721 GGC GGC GCG GCC CAG CGC AAG GTG GCG GCC TTC TTC CGC TCG TGC CTC GAC ATG CGC GAG ATC GAG CGA CTG GGC CCG CGA CCC ATG CTA
 glu val ile glu asp cys gly gly trp asp leu gly gly ala glu glu arg pro gly val ala ala arg trp asp leu asn arg leu leu
 811 GAG GTC ATC GAG GAC TGC GGC GGC TGG GAC CTG GGC GCG GAG GAG CGT CCG GGC GTC GCG CGA TGG GAC CTC AAC CGS CTG CTG
 tyr lys ala gln gly val tyr ser ala ala leu phe ser leu thr val ser leu asp asp arg asn ser ser arg tyr val ile arg
 901 TAC AAG GCG CAG GGC GTG TAC AGC GCC GCC CCG CTC TTC TCG CTC ACG GTC AGC CTG GAC AGG AAC TCC TCG CGC TAC GTC ATC CGC
 ile asp gln asp gly leu thr leu pro glu arg thr leu tyr leu ala gln asp glu ser glu lys val leu ala ala tyr arg val
 991 ATT GAC CAG GAT GGC CTC ACC CTG CCA GAG AGG ACC CTC TAC CTC GCT CAG GAT GAG GAC AGT GAG AAG GTC CTG GCA TAC AGS GTG
 phe met glu arg val leu ser leu leu gly ala asp ala val glu gln lys ala gln glu ile leu gln val glu gln leu ala asn
 1081 TTC ATG GAG CGA GTG-CTC AGC CTC CTC GGT GCA GAC GCT GTG GAA CAG AAG GCC CAA GAG ATC CTG CAA GTG GAG CAG CAG CTG GCC AAC

12/22

352 ile thr val ser glu tyr asp asp leu arg arg asp val ser ser met tyr asn lys val thr leu gly gln leu gln lys ile thr pro
 1171 ATC ACT GTG TCA GAG TAT GAC GAC CTA CGG CGA GAT GTC AGC TCC ATG TAC AAC AAG GTG ACG CTG GGG CAG CTG CAG AAG ATC ACC CCC
 382 his leu arg trp lys trp leu leu asp gln ile phe gln glu asp phe ser glu glu glu val val leu leu ala thr asp tyr met
 1261 CAC TTG CGG TGG AAG TGG CTG CTA GAC CAG ATC TTC CAG GAG GAC TTC TCA GAG GAA GAG GTG GTG CTG GCG ACA GAC TAC ATG
 412 gln gln val ser gln leu ile arg ser thr pro his arg val leu his asn tyr leu val trp arg val val val leu ser glu his
 1351 CAG CAG GTG TCG CAG CTC ATC CGC TCC ACA CCC CAC CGG GTC CTG CAC AAC TAC CTG GTG TGG CGC GTG GTG CTG AGT GAA CAC
 442 leu ser pro pro phe arg glu ala leu his glu leu ala gln glu met glu gly ser asp lys pro gln glu leu ala arg val cys leu
 1441 CTG TCC CCG CCA TTC CGT GAG GCA CTG CAC CAG CAG CTG GCA CAG GAG ATG GAG GGC AGC GAC AAG CCA CAG GAG CTG GCC CGG GTC TGC TTG
 472 gly gln ala asn arg his phe gly met ala leu gly ala leu phe val his glu his phe ser ala ala ser lys ala lys val gln gln
 1531 GGC CAG GCC AAT CGC CAC TTT GGC ATG GCG CTT GGC GGC CTC TTT GTA CAT GAG CAC TTC TCA GCT GCC AGC AAA GCC AAG GTG CAG CAG
 502 leu val glu asp ile lys tyr ile leu gly gln arg leu glu leu asp trp met asp ala glu thr arg ala ala arg ala lys
 1621 CTA GTG GAA GAC ATC AAG TAC ATC CTG GGC CAG CGC CTG GAG CAG CTG GAG TGG ATG GAC GCC GAG ACC AGC GCT GCT CGG GCC AAG
 532 leu gln tyr met met val met val gly tyr pro asp phe leu leu lys pro asp ala val asp lys glu tyr glu phe glu val his glu
 1711 CTC CAG TAC ATG ATG ATG ATG GGC TAC CCG GAC TTC CTG CTG AAA CCC GAT CCT GTG GAC AAG GAG TAT GAG TTT GAG GTC CAT GAG
 562 lys thr tyr phe lys asn ile leu asn ser ile arg phe ser ile gln leu ser val lys lys ile arg gln glu val asp lys ser thr
 1801 AAG ACC TAC TTC AAG AAC ATC TTG AAC AGC ATC CGC TTC ACC CAG CTC TCA CTT AAG AAG ATT CGG CAG GAG CTG GAC AAG TCC ACC
 592 trp leu leu pro pro gln ala leu asn ala tyr tyr leu pro asn lys asn gln met val phe pro ala gly ile leu gln pro thr leu
 1891 TGG CTG CTC CCC CCA CAG CCG CTC AAT GCC TAC TAT CTA CCC AAC AAG AAC CAG ATG GTG TTC CCC GCG GGC ATC CTG CAG CCC ACC CTG
 622 tyr asp pro asp phe pro gln ser leu asn tyr gly gly ile gly thr ile ile gly his glu leu thr his gly tyr asp asp trp gly
 1981 TAC GAC CCT GAC TTC CCA CAG TCT CTC AAC TAC GGG GGC ATC GGC ACC ATC ATT GGA CAT GAG CTG ACC CAC GGC TAC GAC GAC TGG GGG
 652

~~715~~ - 5 (cont'd)

gly gln tyr asp arg ser gly asn leu leu his trp trp thr glu ala ser tyr ser arg phe leu arg lys ala glu cys ile val arg
 2071 GGC CAG TAT CAC CCC TCA GGG AAC CTG CTG CAC TGC TGG ACG GAG GCC TCC TAC ACC CGC TTC CTG CGA AAG GCT CAG TGC ATC GTC CGT 682
 leu tyr asp asn phe thr val tyr asn gln arg val asn gly lys his thr leu gly glu asn ile ala asp met gly gly leu lys leu
 2161 CTC TAT CAC AAC TTC ACT GTC TAC AAC CAG CCG CTG AAC GCG AAA CAC ACG CTT GGG GAG AAC ATC GCA GAT ATG GGC GGC CTC AAG CTG 712
 ala tyr his ala tyr gln lys trp val arg glu his gly pro glu his pro leu pro arg leu lys tyr thr his asp gln leu phe phe
 2251 GCC TAC CAC GCC TAT CAG AAG TGG GTG CCG GAG CAC GCG CCA CAG CAC CCA CTT CCC CGG CTC AAG TAC ACA CAT GAC CAG CTC TTC TTC 742
 ile ala phe ala gln asn trp cys ile lys arg arg ser gln ser ile tyr leu gln val leu thr asp lys his ala pro glu his tyr
 2341 ATT CCC TTT GCC CAG AAC TGG TGC ATC AAC CCG CCG TCG CAG TCC ATC TAC CTG CAG CTG ACT GAC AAG CAT GCC CCT GAG CAC TAC 772
 arg val leu gly ser val ser gln phe glu phe gly arg val leu his cys pro lys val ser pro met asn pro ala his lys cys
 2431 AGG GTG CTG GGC AGT CTG TCC CAG TTT GAG GAG TTT GGC CGG GTT TTA CAC TGT CCA AAG GTC TCA CCC ATG AAC CCT GCC CAC AAG TGT 775
 ser val trp ter
 2521 TCC GTG TGG TGA CCC TGG CTG CCC GGC TGC ACG CCC CCA CTG CCC CCG CAC GAA TCA CCT GCT GGC TAC CGG GGC AGG CAT GCA CCC
 2611 GGT GCC AGC CCC GCT CTG GGC ACC ACC TGC CTT CCA GCC CCT CCA GGA CCC GGT CCC CCT GCT GCC CCT CAC TTC AGG AGG GGC CTG GAG
 2701 CAG GGT GAG GCT GGA CTT TGG GGG GCT GTG AGG GAA ATA TAC TGG GGT CCC CAG ATT CTG CTC TAA GGG GGC CAG ACC CTC TGC CAG GCT
 2791 GGA TTG TAC GGG CCC CAC CTT GGT GGT GCT GCA AGT CTC GTC AAA TAA ATC ACT GCA CTG TTA AAA AAA AAA AAA

13/22

~~FILE~~ - 5 (cont'd)

14/22

Sequence comparison between NEP, NL1, NL2 and NL3

	1	10	20	30	40
NEP-HUM	MG.....	KSESQMDITDINTPKPKKKQ	RWTPLEISLSVLVLL..	LTII...AV	
	*		*	*	
NL1-MOU	MV.....	ERAGWCRKKS	PGFVEYGLMVLLLLLL	GAIVTLG.V..	FYSI.GKQL
	**	* **	* ***	* **	****
NL2-HUM	MV.....	ESAGRAGQKR	PGFLEGGLLLLLLLL	VTAAALVALGV..	YADRRGKQL
	*		* * *	* * *	
NL3-HUM	MEPPYSLTAHYDEFQEVKYVSR	CAGGARGASLPPGFPLGAARSATGARSGLPRWNRREV			

	50	60	70
NEP-HUM	TMIA.....	LYATYDD.....	GICKSSDCIKSAARLIQ.NMDATT
	*	*	** * ****
NL1-MOU	PLLTSL....	LHFSWDERTVVKR...ALRDSSLKSDICTTPSCVIAAARILE.NMDQSR	
	* * *	* * * * *	* * * * * * * * *
NL2-HUM	PRLASR....	LCFLQEERTFVKRKRPGIPEAQEVSEVCTTPGCVIAAARILQ.NMDPTT	
	*	**	* * * * *
NL3-HUM	CLLSGLVFAAGLCAILAAMLALKYLGPVAAGGGACPEGC	PERKAFARAARFLAANLDASI	

	80	90	100	110	120	130
NEP-HUM	EPCTDFFKYACGGWLKRN	VIPETSSRYGNFDILRDELEVVLKDV	LQEPKTEDIVAVQ.KA			
	** *	*****	*****	*****	** *	**
NL1-MOU	NPCENFYQYACGGWLRHH	VIPETNSRYSVFDILRDELEVILKGVLEDSTSQHRPAVE.KA				
	** *	*****	*****	*****	*****	**
NL2-HUM	EPCDDFYQFACGGWLRHH	VIPETNSRYSIFDVLRLDELEVILKAVLENSTAKDRPAVE.KA				
	** ***	*****	*	* * *	*	*
NL3-HUM	DPCQDFYSFACGGWLRRAI	PDDKLTGYGTIAAIGE	QNEERLRRLRLARPGGGPGGAAQRKV			

	140	150	160	170	180	190
NEP-HUM	KALYRSCINESAIDSRGGEPL	LKLLPDIYGWPVATENWEQKYGASWTA	EKAIAQLNSKYG			
	* *****	* ****	* *****	* *	* * *	* ****
NL1-MOU	KTLYRSCMNQSVIEKRDSEPL	LSVLKMGVGGWPVAMDKWNETMGLKWELERQLAVLNSQFN				
	*****	* ****	* *****	* * *	* * *	*****
NL2-HUM	RTLYRSCMNQSVIEKRG	SQPLLDILEVVGWPVAMDRWNETVGLEWELERQLALMNSQFN				
	* ***	* * *	* * *	* * *	* * *	*
NL3-HUM	RAFFRSLDMREIERL	GPRLPMLEVI	EDCGGWDLGGA	EERPGVAARWDLNRLLYKAQGVYS		

	200	210	220	230	240	250
NEP-HUM	KKVLINL	FGTDDKNSVNHVIHIDQ	PRLGLPSRDYYECTGIYKEACTAYVDFMISVARLI			
	*** **	* * *	* * * * *	* * * * *	* * *	* * *
NL1-MOU	RRVLIDLFIWNDDQNSSRH	VIYIDQPTLGMP	SREYYFQEDNNHKVRKAYLEFMTSVATML			
	*****	*****	*****	* * *	* * *	* * *
NL2-HUM	RRVLIDLFIWNDDQNSSRH	IIYIDQPTLGMP	SREYYFN	GGSNRKVREAYLQFMVSVATLL		
	* *	* * *	* * *	* * *	*	*
NL3-HUM	AAALFSLTVSLDDR	NSSRYVIRIDQDGLTL	PERTLYLAQDEDESE..KVLAAYRVFMERVL			

15/22


	260	270	280	290	300	310
NEP-HUM	RQEERLPIDENQLALEMNKVMELEKEIANATAKP..EDRNDPMLLYNKMTLAQIQNNFSL					
	* * *	* * *	* * *	* * *	* * *	* * *
NL1-MOU	RKDQNL SKESAMVREEMA EVLELETHLANATVPQ..EKRHDVTALYHRMDLMELQERFGL					
	* * *	* * *	* * *	* * *	* * *	* * *
NL2-HUM	REDANLPRDSCLVQEDMVQVLELETQLAKATVPQ..EERHDVIALYHRMGLEELQSQFGL					
	* * *	* * *	* * *	* * *	* * *	* * *
NL3-HUM	SL...LGADAV..EQKAQEILQVEQQLANITVSEYDDLRRDVSSMYNKVTILGQLQKITP.					

	320	330	340	350	360	370
NEP-HUM	EINGKPF SWLNFTNEIMSTVNISITNEEDVVVYAPEYLT KLKPI LTKYSARDLQNLMSWR					
	* * *	* * *	* * *	* * *	* * *	* * *
NL1-MOU	...KGFNWTLFIQNVLSSEVELEFPDEEVVVGIPYLENLEDIIDSYSARTMQNYLVWR					
	* * *	* * *	* * *	* * *	* * *	* * *
NL2-HUM	...KGFNWTLFIQTVLSSVKIKLLPDEEVVVGIPYLQNLNIIDTYSARTIQNYLVWR					
	* * *	* * *	* * *	* * *	* * *	* * *
NL3-HUMHLRWKLLDQIF...QEDFSEEEVVLLATDYMQQVSQLIRSTPHRVLHNYLVWR					

	380	390	400	410	420	430
NEP-HUM	FIMDLVSSLSR TYKESRNAFRKALYGT TSETATWRR CANVNGNMENAVGRLYVEAA FAG					
	* * *	* * *	* * *	* * *	* * *	* * *
NL1-MOU	LVLDRIGSLSQR FKEARVDYRKALYGT TVEEVRWREC VSYVNSNMESAVGSLYIKRAF SK					
	* * *	* * *	* * *	* * *	* * *	* * *
NL2-HUM	LVLDRIGSLSQR FKDTRVNYRKALFG TMVEEVRWREC VGYVNSNMENAVGSLYVREAF PG					
	* * *	* * *	* * *	* * *	* * *	* * *
NL3-HUM	VVVVLSEHLSPPFREALHELAQEMEGSDKPQELARVCLGQANRHF GMALGALFVHEHFSA					

	440	450	460	470	480	490
NEP-HUM	ESKHVVEDLIAQIREVFIQTLDDLTWMDAETKKRAEEKALAIKERIGYPDDIVSNDNK.L					
	* * *	* * *	* * *	* * *	* * *	* * *
NL1-MOU	DSKSTVRELIEKIRSVFVDNLDELNWMDEESKKKAQEKAMNIREQIGYPDYILEDNNKHL					
	* * *	* * *	* * *	* * *	* * *	* * *
NL2-HUM	DSKSMVRELIDKVRTVFVETLDELGWMDEESKKKAQEKAMSIREQIGHPDYILEEMNRRL					
	* * *	* * *	* * *	* * *	* * *	* * *
NL3-HUM	ASKAKVQQLVEDIKYILGQRLEELDWMDAETRAAARAKLQYMMVMVGYPDFLLKPDA..V					

	500	510	520	530	540	550
NEP-HUM	NNEYLELNYKEDEYFENIIQNLKFSQSKQLKKLREKVDKDEWISGAAVVNAFYSSGRNQI					
	* * *	* * *	* * *	* * *	* * *	* * *
NL1-MOU	DEEYSSLT FYEDLYFENGLQNLKNN AQRSLKKLREKVDQNLWII GAAVVNAFYSPNRNQI					
	* * *	* * *	* * *	* * *	* * *	* * *
NL2-HUM	DEEYSNLNFS EDLYFENSLQNLKVG AQRSLRKLREKVDPNLWII GAAVVNAFYSPNRNQI					
	* * *	* * *	* * *	* * *	* * *	* * *
NL3-HUM	DKE.YEFVHEKTYFKNILNSIRFSIQLSVKKIRQEVDKSTWLLPPQALNAYYLPKNQM					

 (cont'd)

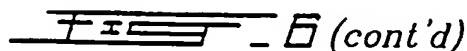
16/22

	560	570	580	590	600	610
NEP-HUM	VFPAGILQPPFFSAQQSN	SLNYGGIGMVIGHEITHG	FDDNGRNFNKGDLVDWWT	QQSAS		
	*****	*	*****	*****	*	*****
NL1-MOU	VFPAGILQPPFFSKDQPQ	SLNFGGIGMVIGHEITHG	FDDNGRNFNKGDLVDWWT	QQSAS		
	*****	*****	*****	*****	*****	*****
NL2-HUM	VFPAGILQPPFFSKEQPQ	ALNFGGIGMVIGHEITHG	FDDNGRNFNKGDLVDWWT	QQSAS		
	*****	*****	*****	*****	*****	*****
NL3-HUM	VFPAGILQPTLYDPDFPQ	SLNYGGIGTII GHELTHG	YDDWGGQYDRSGNLLHW	TEASYS		

	620	630	640	650	660	670
NEP-HUM	NFKEQSQCMVYQYGNFSW	DLAGGQHLNGINTLGENI	ADNNGGLGQAYRAYQNYI	KKNGEEK		
	*	*****	*****	*****	*****	*
NL1-MOU	HFQQQSQCMYQYGNFSW	ELADNQNNGFSTLGENI	ADNNGGVRQAYKAYLRWL	ADGGKDQ		
	**	*****	*****	*****	*****	*****
NL2-HUM	HFREQSECMYQYGNYSW	DLADEQNNGFNTLGENI	ADNNGGVRQAYKAYLKWM	AEGGKDQ		
	*	*****	*****	*****	*****	*****
NL3-HUM	RFLRKAECIVRLYDNFT	...VYNQRVNGKHTLGENI	ADMGGLKLAYHAYQKWV	REHGPEH		

	680	690	700	710	720	730
NEP-HUM	LLPGLDLNHKQLFFLNFA	QVWCGTYRPEYAVNSIKT	DVHSPGNFRIIGTLQNSA	EFSEAF		
	*****	*****	*****	*****	*****	*****
NL1-MOU	RLPGLNLTYAQLFFINYA	QVWCGSYRPEFAVQSIKT	DVHSPLKYRVLGSLQNL	PGFSEAF		
	*****	*****	*****	*****	*****	*****
NL2-HUM	QLPGLDLTHEQLFFINYA	QVWCGSYRPEFAIQSIKT	DVHSPLKYRVLGSLQNL	AFAFADF		
	*****	*****	*****	*****	*****	*****
NL3-HUM	PLPRLKYTHDQLFFIAFA	QNWCIKRRSQSIYLQVLT	DKHAPEHYRVLGSVSQF	EEFGRVL		

	740	750
NEP-HUM	HCRKNSYMNPEKKCRVW	
	**	*****
NL1-MOU	HCPRGSPMHPMKRCRIW	
	**	*****
NL2-HUM	HCARGTPMHPKERCVRW	
	**	*****
NL3-HUM	HCPKVSPMNPAAHKCSVW	

 (cont'd)

17/22

NL1 in the TESTIS

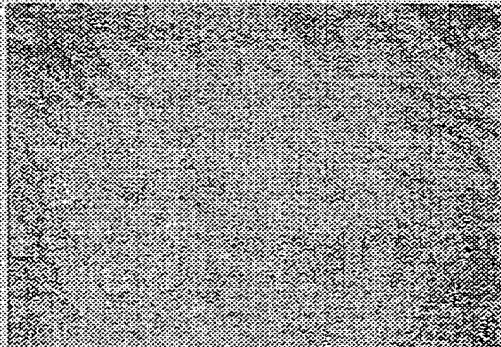
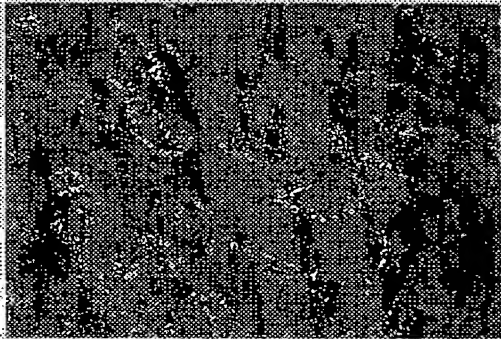
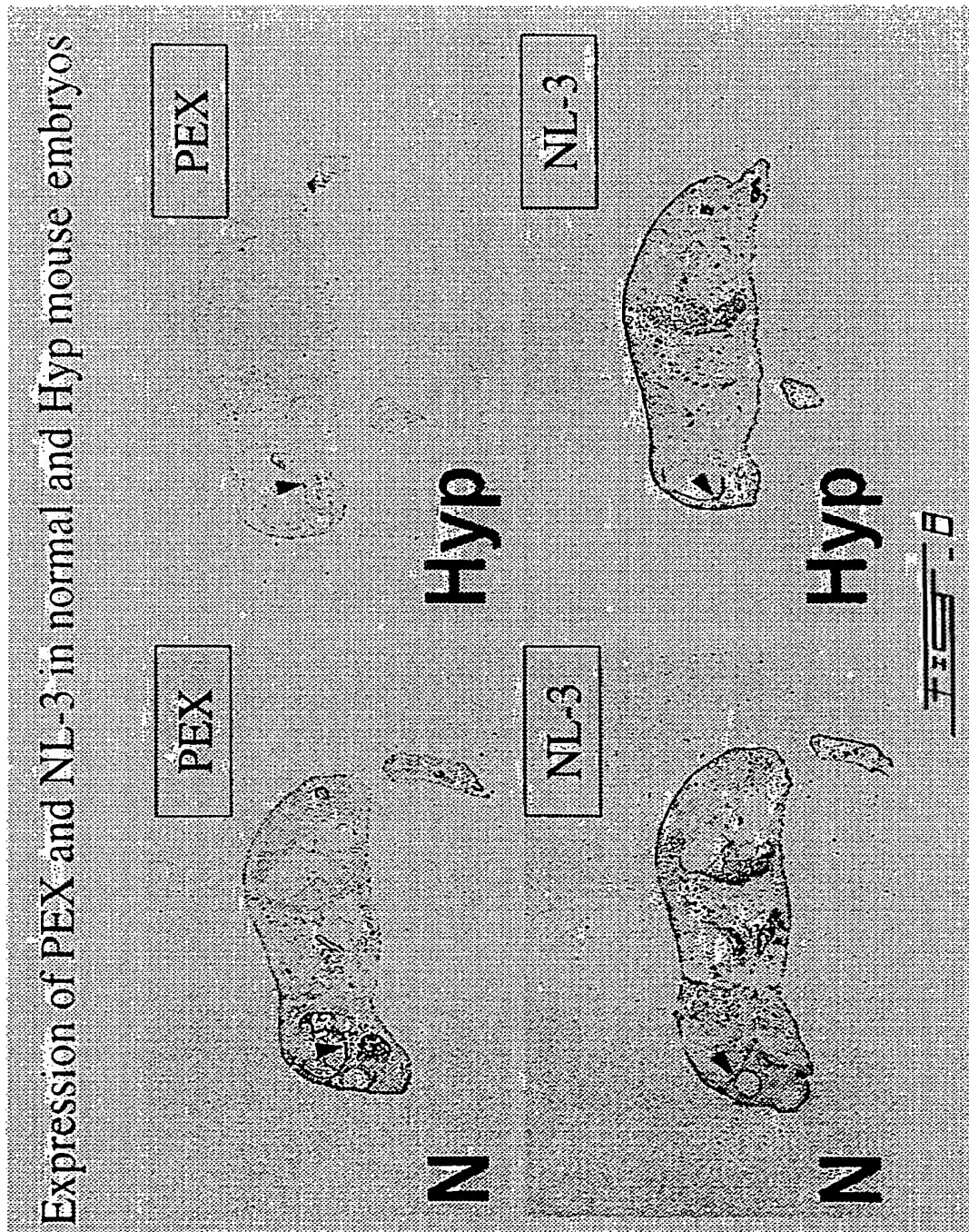


Fig. 7

BEST AVAILABLE COPY

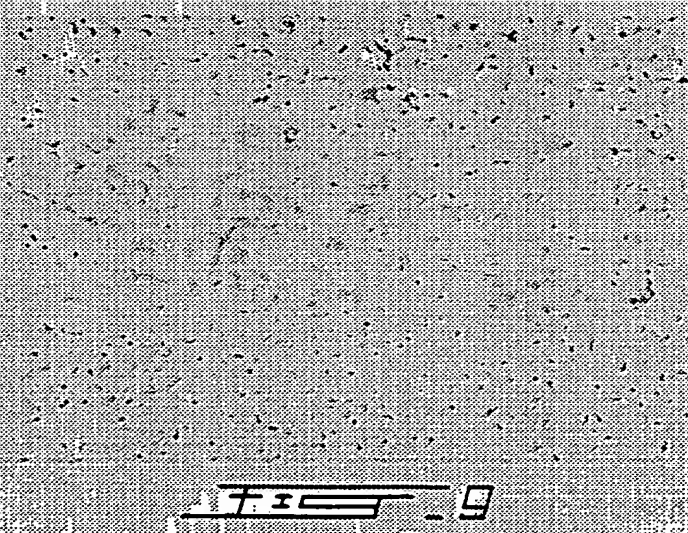
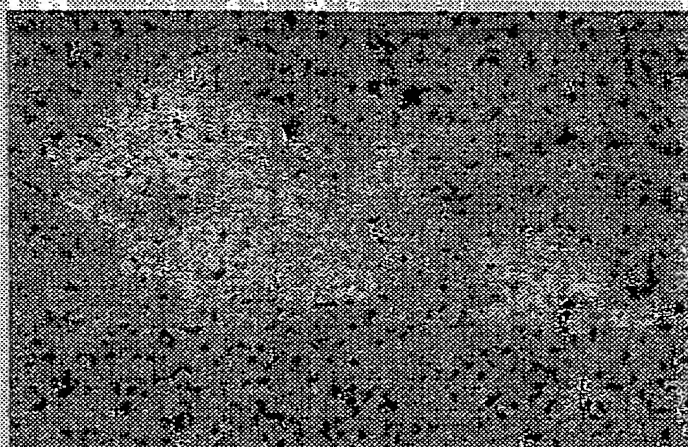
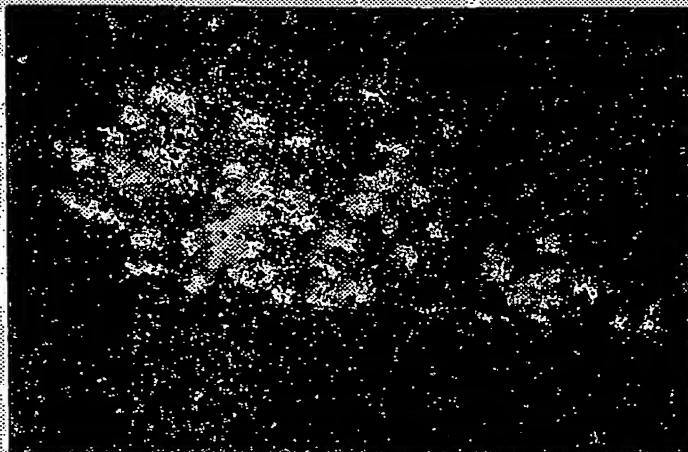
18/22



BEST AVAILABLE COPY

19/22

NL3 in the BRAIN

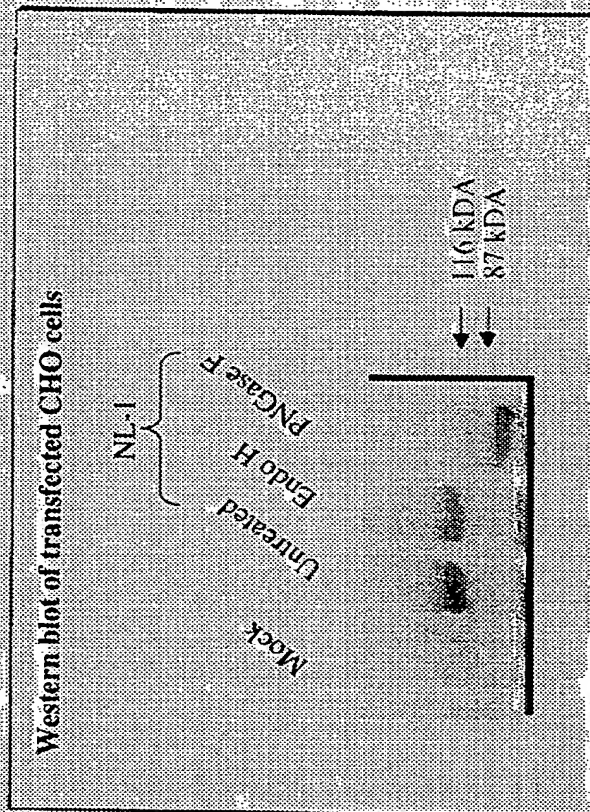
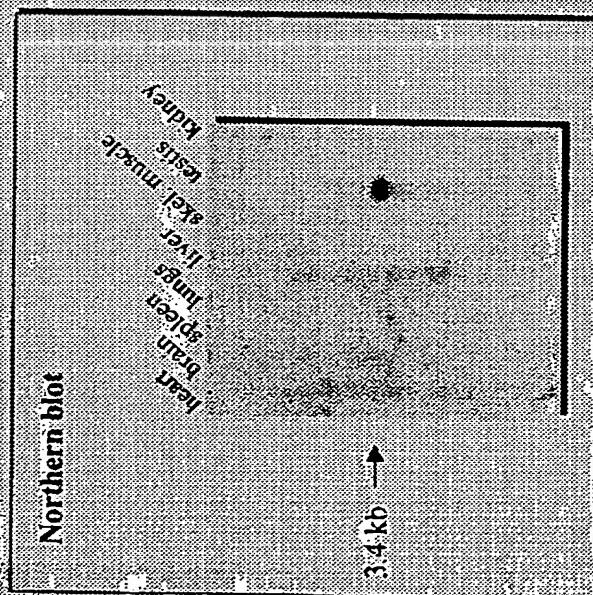
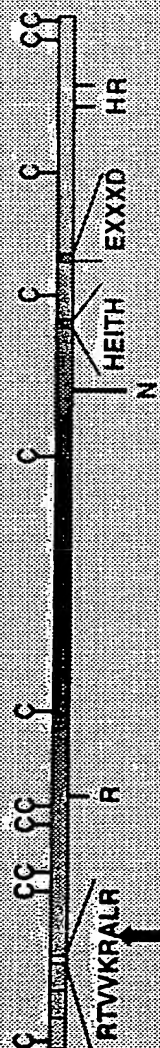


BEST AVAILABLE COPY



20/22

Structure and expression of NL-1



01-11-11

BEST AVAILABLE COPY

21/22

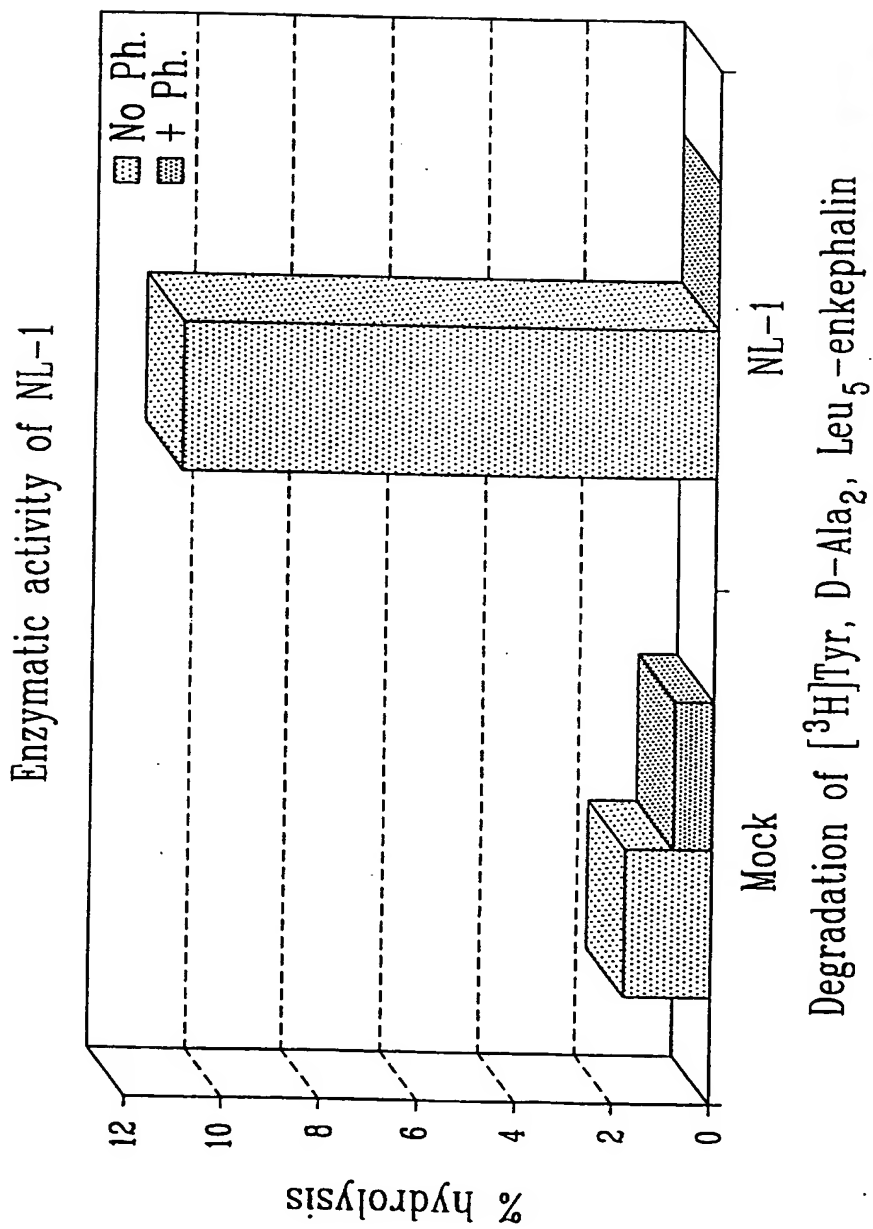


FIG. 11

BEST AVAILABLE COPY

22/22

BEST AVAILABLE COPY

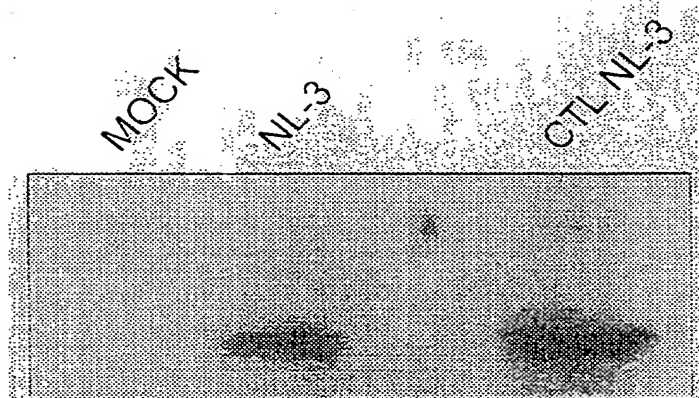


FIG. 12